

Systematic Analysis of Lysine Acetyltransferases

Dissertation

zur Erlangung des akademischen Grades Dr. rer. nat.
vorgelegt der Fakultät für Biologie der
Ludwig-Maximilians-Universität München

Christian Feller

München 2014

1. Gutachter: Prof. Dr. Peter B. Becker

2. Gutachter: Prof. Dr. Dirk Eick

Dissertation eingereicht am: 16.09.2014

Mündliche Prüfung am: 28.10.2014

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

München, den 15. September 2014

.....

(Christian Feller)

Erklärung

Hiermit erkläre ich, dass die vorgelegte Arbeit an der LMU von Herrn Prof. Dr. Peter Becker betreut wurde.

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in Teilen einer anderen Prüfungskommission vorgelegt worden ist. Weiterhin habe ich weder an einem anderen Ort eine Promotion angestrebt noch angemeldet, noch versucht eine Doktorprüfung abzulegen.

Die eigenen Leistungen für die in dieser kumulativen Dissertation enthaltenen Manuskripte sind in den Kapiteln 3.1.1, 3.2.1, 3.3.1 und 3.4.1 aufgelistet.

München, den 15. September 2014

.....

(Christian Feller)

For my parents

Table of contents

1	SUMMARY	11
2	INTRODUCTION	15
2.1	The nucleosome is the basic repeat unit of chromatin	16
2.2	Higher-order chromatin structures	18
2.3	Post-translational histone modifications	20
2.3.1	Lysine acetylation	21
2.3.2	Lysine methylation	30
2.4	Dosage compensation in the fruit fly <i>Drosophila melanogaster</i>	33
3	RESULTS AND DISCUSSION	36
3.1	The activation potential of MOF is constrained for dosage compensation	37
3.1.1	Summary, significance and own contribution	38
3.1.2	Published manuscript	40
3.1.3	Supplementary data and figures	53
3.2	Dosage compensation and the global re-balancing of aneuploid genomes	72
3.2.1	Summary and own contribution	73
3.2.2	Published review article	74
3.3	The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset	83
3.3.1	Summary, significance and own contribution	84
3.3.2	Published manuscript	86
3.3.3	Supplementary data and figures	101
3.4	Global and specific responses of the histone acetylome to systematic perturbation	114
3.4.1	Summary, significance and own contribution	115
3.4.2	Submitted manuscript	117
3.4.3	Supplementary data, tables and figures	151
3.4.4	Discussion on substrate specificity of lysine acetyltransferases using the examples of HAT1 and HBO1	203
4	GENERAL CONCLUSIONS AND OUTLOOK	208
5	REFERENCES	213
6	ACKNOWLEDGEMENTS	236
7	LIST OF ABBREVIATIONS	238
8	CURRICULUM VITAE	241

1 SUMMARY

Eukaryotic genomes are packed in chromatin that comprises an ever repeating succession of nucleosomes with DNA wrapped around octamers of histone proteins. Dynamic regulation of chromatin structure enables controllable access to the underlying DNA and hence is crucial to all nuclear processes, including DNA transcription, replication and repair. Many interconnected mechanisms are in place to regulate chromatin structure. These include chemical modifications of histone proteins, such as lysine acetylation and methylation, which directly alter the properties of nucleosomes to form repressive structures or install signaling marks for dedicated effector proteins.

The original research described in this cumulative thesis is contained in two published articles and one submitted manuscript centering on histone lysine acetylation in *Drosophila melanogaster*. In the first two articles, we characterise the functions of two multi-protein complexes containing the lysine acetyltransferase MOF (males absent on the first). If incorporated in the male-specific lethal dosage compensation complex (MSL-DCC) MOF acetylates lysine 16 on histone H4 (H4.K16ac) on gene bodies on the male X chromosome, which is critical for the two-fold transcriptional stimulation of its target genes. The underlying process, dosage compensation, serves to adjust gene expression levels between the single male and the two female X chromosomes. Combining genome-wide mapping and transcriptome studies with the analysis of defined reporter loci in transgenic flies and cell lines, we provided evidence that MOF-mediated H4.K16ac has an inherent strong transcriptional activation potential, which is, however, constrained to the physiological two-fold range in the context of fly dosage compensation. In contrast, MOF as a component of the non-specific lethal (NSL) complex binds promoters of active housekeeping genes located on all chromosomes in male and female flies. However, transcriptional activation through the NSL complex is found only at a subset of these binding sites. These NSL-regulated genes are enriched for a specific core promoter sequence and depleted for the insulator proteins CP190 and BEAF as well as the heterochromatin protein HP1c. In summary, these studies describe the context-specific localization and transcriptional activation modes of the acetyltransferase MOF through its incorporation into two distinct multi-protein complexes.

Histones have acquired impressive patterns of acetylation sites, where the individual abundance and the context of co-occurrence with other marks potentially confers very different functions. However, the current methodologies are too limited to systematically evaluate changes in rare and combinatorial modification motifs. Moreover, how different enzymes contribute to complex motifs is only poorly understood. In the third study, we generated a catalogue of histone acetylation and methylation motifs that describe the changes in response to ablation of every known or suspected lysine acetyltransferase and deacetylase. To achieve this goal, we first optimised liquid chromatography mass spectrometry workflows that enabled highly accurate and precise quantification of combinatorial histone

modification motifs. The comprehensiveness of the dataset not only allowed us to recognise class-specific properties to these enzyme families but also to describe the histone modification system as an interconnected, flexible network that compensates the loss of an individual component. We describe the specific responses of the histone acetylome upon ablation of each acetyltransferase. To our surprise we also observed that depletion of almost every acetyltransferase triggered a systemic response that effectively maintained global histone acetylation levels. Finally, we documented that dosage compensation is not restricted to the change of the single H4.K16ac mark, but accompanied by a specific re-distribution of acetylation and methylation motifs. In summary, this study provides further evidence that chromatin pathways are highly interconnected and highlights the necessity to study the function of each individual component in the context of the system.

ZUSAMMENFASSUNG

Eukaryotische Genome sind in einer Chromatinstruktur eingebettet. Chromatin besteht in seiner einfachsten Organisationsform aus einer Aneinanderreihung von Nukleosomen, in denen DNA spulenförmig um Oktamere aus Histonproteinen gewunden ist. Die Chromatinstruktur wird sehr dynamisch reguliert. Dies ist wichtig für Prozesse wie die Transkription, Replikation und DNA Reparatur, bei denen Faktoren kontrollierten Zugriff zur DNA benötigen. Im Zellkern existieren viele ineinandergreifende Mechanismen, um die Chromatinstruktur für die jeweiligen Prozesse zu regulieren. Dazu gehören chemische Modifikationen an den Histonproteinen, wie zum Beispiel Lysin-Acetylierung und Lysin-Methylierung. Diese Modifikationen haben einerseits das Potenzial, direkt das Gleichgewicht zwischen kompakten und zugänglichen Chromatinorganisationsformen zu beeinflussen. Zum anderen wirken Histonmodifikationen als Signale für Effektorproteine, welche wiederum Änderungen in der Chromatinstruktur herbeiführen oder direkt biologische Prozesse wie die Transkription in Gang setzen können.

Diese kumulative Dissertation enthält zwei publizierte und einen eingereichten Originalartikel, die sich mit dem Thema der Histon-Acetylierung in der Fruchtfliege *Drosophila melanogaster* auseinandersetzen. In den ersten beiden Artikeln charakterisieren und vergleichen wir die Funktionen von zwei Multiproteinkomplexen, welche die Acetyltransferase MOF beinhalten. MOF innerhalb des „Dosis-Kompensations-Komplexes“ acetyliert Lysin 16 auf Histone H4 (H4.K16ac) entlang der Gene auf dem männlichen X-Chromosom. Diese Acetylierung ist wichtig für den Prozess der Dosiskompensation, bei dem durch die zweifache Stimulation der Transkription von Genen auf dem einzelnen männlichen X-Chromosom die „Gen-Dosis“ gegenüber den Genen auf den zwei weiblichen X-Chromosomen angepasst wird. In unserer Studie haben wir genomweite Lokalisations- und Transkriptionskartierungsmethoden mit der Analyse von definierten Reportergenen in transgenen Fliegen und Zellsystemen verbunden. Dies erlaubte uns festzustellen, dass zwar die MOF-vermittelte H4.K16-Acetylierung ein starkes Potenzial hat, die Transkription zu stimulieren, dieses Potenzial jedoch im Kontext der Dosiskompensation auf eine feinjustierte zweifache Erhöhung gedrosselt ist. Im Gegensatz dazu fungiert MOF innerhalb des NSL-Komplexes (*non-specific lethal*) als Promotertypischer Ko-Aktivator. Der NSL-Komplex bindet Promotoren von „Haushaltsgenen“ entlang aller Chromosomen in männlichen sowie in weiblichen Zellen. Unsere detaillierte genombiologische-bioinformatische Analyse zeigte allerdings, dass nur ein Teil der gebundenen Gene auch vom NSL-Komplex reguliert werden. Diese NSL-regulierten Gene enthalten eine spezifische Promotersequenz-Signatur und werden im Gegensatz zu den anderen gebundenen aber nicht-regulierten Genen tendenziell nicht von den Insulatorproteinen CP190 und BEAF sowie vom Heterochromatinprotein HP1c gebunden. Zusammenfassend zeigen diese beiden Studien, dass die Lokalisationen und Funktionen von MOF stark von den assoziierten Multiproteinkomplexen geprägt werden.

Histone zeigen eine hohe und beeindruckende Anzahl von Modifikationsmustern. Die individuelle Häufigkeit sowie das gemeinsame Auftreten mit anderen Modifikationen vermitteln höchstwahrscheinlich sehr verschiedene biologische Funktionen. Die technischen Möglichkeiten, um niedrig-abundante sowie spezielle Kombinationen von Modifikationsmustern zu detektieren und deren Änderungen zu quantifizieren, sind zurzeit noch sehr begrenzt. Des Weiteren ist es noch unzureichend verstanden, wie verschiedene Enzyme zu komplexen Modifikationsmotiven beitragen. In der dritten Studie haben wir umfassend katalogisiert, wie sich einfache und komplexe Acetylierungs- und Methylierungsmotive auf Histon-Proteinen ändern, nachdem wir systematisch jede bekannte oder vermutete Acetyltransferase und Deacetylase aus *Drosophila*-Zellen entfernt haben. Um dies zu erreichen, war es zuerst nötig, spezielle Protokolle der Flüssigchromatographie-Massenspektrometrie (LC-MS) zu optimieren, welche uns erlaubten, Änderungen in kombinatorischen Modifikationsmustern akkurat und präzise zu quantifizieren. Der große Umfang dieses Kataloges in zwei Dimensionen – alle bekannten und vermuteten Enzyme gegenüber vielen Modifikationsmustern – erlaubte uns nicht nur, enzymklassenspezifische Eigenschaften zu erkennen und zu beschreiben, sondern auch das Histon-Modifikations-System als ein höchstverwobenes und flexibles Netzwerk zu dokumentieren, in dem der Verlust von einzelnen Komponenten systemisch kompensiert wird. Der überraschendste Befund war die Feststellung, dass nach dem einzelnen Entfernen fast jeder Acetyltransferase die Gesamtbilanz aus acetylierten und nicht-acetylierten Histonen konstant war. Schließlich untersuchten wir genauer, welche systemischen Auswirkungen der Prozess der Dosiskompensation auf die Histonmodifikationslandschaft mit sich bringt. Hier stellten wir fest, dass Dosiskompensation nicht nur auf dem bekannten Acetylierungssignal H4.K16ac beruht, sondern dass sich systemweit auch andere Acetylierungs- und Methylierungsmotive umverteilen. Die Ergebnisse dieser Studie erbringen weitere Hinweise, dass Chromatinregulationsmechanismen komplexe, miteinander verbundene Signalwege darstellen, bei denen die Funktion einzelner Komponenten im Kontext des Systems studiert werden sollte.

2 INTRODUCTION

The evolution of the eukaryotic cell is accompanied by impressive innovations that enable the development of highly complex multi-cellular organisms and ensure adequate responses to internal and external stimuli. One fascinating invention was the organisation of eukaryotic genomes into chromatin. Chromatin is a dynamic structure that consists of an ever repeating succession of a single fundamental ‘building block’, the nucleosome. In the nucleosome, DNA is wrapped around an octamer of four histone proteins, namely H2A, H2B, H3 and H4. Chromatin exists in a continuum of states that allow (euchromatin) or permit (heterochromatin) access to the underlying DNA. Many interconnected mechanisms determine the structure and function of chromatin and thereby regulate most biological processes that utilise DNA. These mechanisms include i) post-translational, chemical modifications at histones, ii) ATP-dependent chromatin remodelling, iii) the replacement of canonical histones by histone variants, iv) association of linker histones and constitutive non-histone proteins, v) DNA methylation, vi) non-coding RNAs and vii) the organisation of chromatin fibres within the three-dimensional nuclear architecture. The concerted action of these mechanisms allows the implementation of specific genetic programs.

In this introduction, I will primarily focus on describing post-translational histone modifications (PTMs) and their significance in regulating chromatin structure and transcription. After introducing the nucleosome and higher-order chromatin structure, I will describe post-translational modifications through lysine acetylation on histones, which is the main topic of my PhD studies. In addition, I will briefly discuss lysine methylation at histones because there is well-documented cross-talk between the two types of modifications. Finally, I will introduce the phenomenon of *Drosophila* dosage compensation that provides an illustrative example of how a single histone acetylation site contributes to the transcriptional upregulation of an entire chromosome.

2.1 The nucleosome is the basic repeat unit of chromatin

The term chromatin was coined by Walther Flemming in the 1890s describing nuclear structures that stain strongly with basophilic dyes (Flemming 1882). Around the same time, analysing the chemical composition of the cell nucleus, Friedrich Miescher identified nucleic acid and Albrecht Kossel characterised a protein component that he termed *Histon* (Miescher 1871, Kossel 1884). However, it was not before the 1970s when experiments on endonuclease-mediated chromatin digestion, histone cross-linking and electron microscopy of chromatin particles revealed the nucleosome as the fundamental repeat unit of chromatin (Hewish and Burgoyne, 1973; Kornberg, 1974; Kornberg and Thomas, 1974; Olins and Olins, 1974; Oudet et al., 1975).

Another 20 years later, Richmond and colleagues solved the crystal structure of a nucleosome at 2.8 Å resolution (Figure 1, (Luger et al., 1997)). It reveals that 147 base pairs (bp) of DNA are wrapped in 1.65 left-handed (negative) superhelical turns around an octamer of the canonical histones H2A, H2B, H3 and H4. Positive charges of the histone proteins contact the negatively charged phosphate backbone of DNA every ~10.4 bp, establishing 14 relatively weak histone-DNA contacts. The histone octamer is a modular assembly formed by association of two H3-H4 dimers to form an (H3-H4)₂ tetramer followed by binding of two H2A-H2B dimers on either side of the tetramer. Each histone consists of a globular core domain and N- and C-terminal domains ('tails') that protrude from the core particle. The globular core constitutes the characteristic histone fold domain, which is comprised of three α -helices separated by two loops. Interactions between the histone folds of H3 stabilise the H3-H4 pairs whereas the H2A-H2B pairs interact with the tetramer through a homologous 4-helix bundle between the H2B and H4 folds. Further contacts between the H2A docking domain towards H3-H4 and among the H2A loop 1 regions of the H2A-H2B dimers stabilise the nucleosome. Of particular importance is the interaction between an acidic patch at the H2A-H2B dimer with the basic patch at the H4 tail because it is a critical target for the regulation of nucleosome stability and chromatin structure through the action of chromatin remodelling enzymes and histone acetylation (discussed in detail in 2.3.1).

The nucleosome is not a static particle but rather shows transient DNA breathing and 'open nucleosome states'. Recent studies applying fluorescence resonance energy transfer (FRET), high-speed atomic force microscopy (HS-AFM) and small-angle X-ray scattering (SAXS) provided evidence for a dynamic nucleosome where DNA reversibly dissociates, H2A-H2B dimers partially disrupt from the (H3-H4)₂ tetramer while still associated to DNA and nucleosomes undergo spontaneous sliding and even complete dissociation in the absence of ATP-dependent remodelling (Mangenot et al., 2002; Li et al., 2005; Zlatanova et al., 2009; Andrews and Luger, 2011; Bohm et al., 2011; Miyagi et al., 2011; Luger et al., 2012). These findings have strong implications for concepts how transcription factors access DNA within the context of chromatin.

A further diversification of nucleosome structure and its stability arises from replacing canonical histones by histone variants (Hake and Allis, 2006; Talbert and Henikoff, 2010; Bonisch and Hake, 2012). For example, the centromere-specific histone variant CID (CENP-A in humans) was proposed to form stable tetramer structures that organise only 120 bp of DNA in a right-handed (positive) manner (Dalal et al., 2007; Furuyama and Henikoff, 2009). While the histone variant macroH2A increases nucleosome stability, H2A.Bbd containing nucleosomes do not form octamers under high-salt conditions and show faster FRAP mobility (reviewed in (Bonisch and Hake, 2012)).

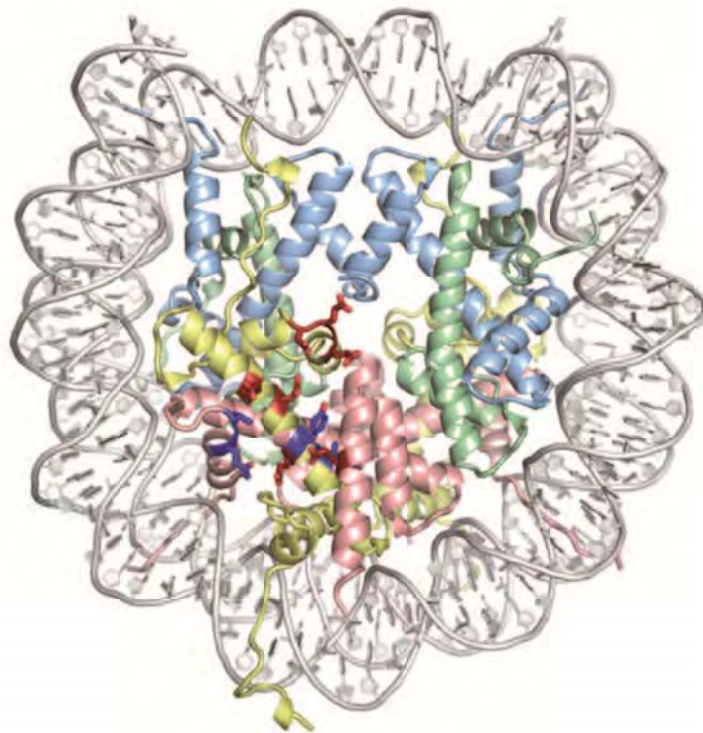


Figure 1. The nucleosome structure.

The structure of the nucleosome is visualised down the superhelical axis of the DNA. DNA is shown in grey and histones H2A, H2B, H3 and H4 in yellow, red, light blue and green. Selected acidic residues from H2A and H2B that are important for the interaction with the histone H4 tail are shown in dark red. Adapted from (Luger et al., 2012).

2.2 Higher-order chromatin structures

Nucleosomes are arranged on DNA like ‘beads on a string’ forming the primary structure of chromatin, called the ‘10 nm fibre’. The nucleosomes are separated by linker DNA, which show variable length distributions typically between 10 bp and 50 bp in a context- and species-specific manner (Clapier and Cairns, 2009). Linker histones such as H1 and H5 bind to DNA at the entry and exit sites of nucleosomes and protect 20 bp of linker DNA against micrococcal nuclease. Although linker histones are not part of the nucleosome core particle they contribute to the secondary structure of chromatin, such as the 30 nm fibre described *in vitro*.

Linear arrays of nucleosomes (primary structure) are proposed to fold into three-dimensional assemblies of higher-order structures (e.g. 30 nm fibre). Two competing models of the 30 nm fibre have been put forward based on *in vitro* experiments on reconstituted chromatin fibres and the crystal structure of a tetranucleosome particle (Figure 2). In the ‘one-start’ solenoid model, neighbouring nucleosomes interact with each other and follow a helical trajectory with six to eight nucleosomes per turn (Finch and Klug, 1976; Widom and Klug, 1985; Robinson et al., 2006; Routh et al., 2008; Kruithof et al., 2009). The two-start zigzag model proposes that two rows of nucleosomes form a two-start helix with interactions between alternating nucleosomes (Dorigo et al., 2004; Schalch et al., 2005; Song et al., 2014). It should be stressed that parameters such as linker length, the size of the nucleosomal array and whether H1, H5 or no linker histone was used strongly favour the one or the other model (Routh et al., 2008; Grigoryev et al., 2009).

In contrast to the above mentioned *in vitro* studies, whether higher order chromatin structures exist *in vivo* is controversially discussed (Horowitz et al., 1990; Tremethick, 2007; Maeshima et al., 2010; Nishino et al., 2012). Evidence supporting the existence of 30 nm fibres are currently limited to highly specialised cell types, including transcriptionally silent chicken erythrocytes and starfish spermatozoa (Woodcock, 1994; Horowitz et al., 1997). In contrast, there is yet no evidence for higher order chromatin structures in interphase and metaphase chromosomes examining different species and cell types, including mitotic HeLa cells (Tremethick, 2007; Maeshima et al., 2010; Nishino et al., 2012). According to the recently proposed ‘polymer melt’ model, the high concentration of nucleosomes in the nucleus preferentially drives inter-fibre interactions so that individual fibres strongly interdigitate. However, the polymer melt model also predicts that 30 nm fibres form transiently during transitions between different chromatin condensation states such as the ones that may occur upon transcriptional activation (Maeshima et al., 2010). Together, these studies suggest that there is no uniform secondary structure of chromatin in the cell. Instead, highly dynamic global and local transitions in chromatin structure provide an excellent window of opportunity for regulatory mechanisms (Tremethick, 2007; Maeshima et al., 2010; Fussner et al., 2011).

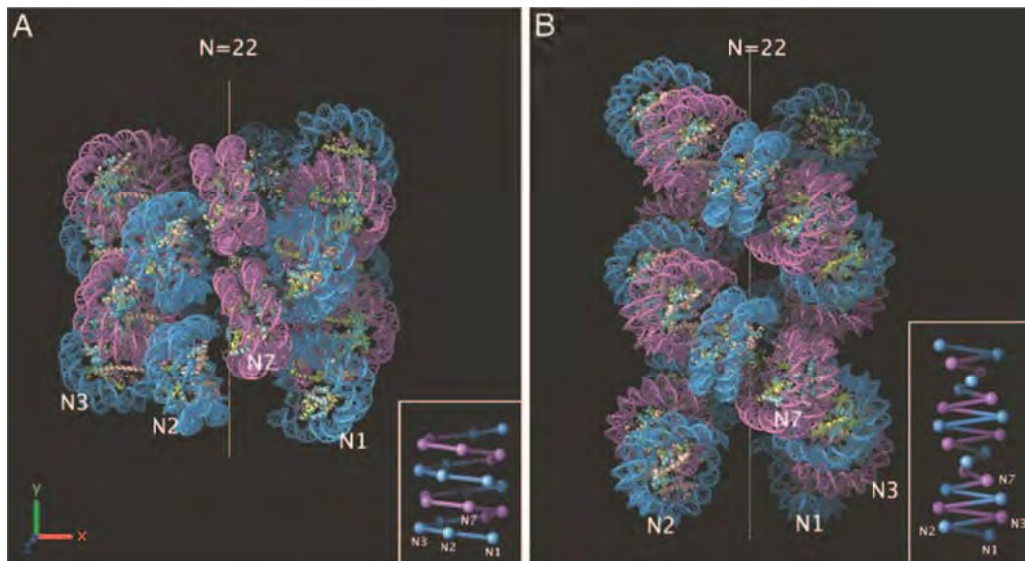


Figure 2. One-start and two-start helix chromatin fibre models.

Idealised nucleosomal fibre models of the interdigitated one-start helix model (A) and the two-start zigzag helical model (B). Alternating nucleosomes are shown in blue and magenta. Nucleosomes 1, 2, 3 and 7 are indicated. Insert shows proposed connectivity of DNA. Figure from (Robinson et al., 2006).

The structure of chromatin is regulated primarily through the action of ATP-dependent chromatin remodelling enzymes. Chromatin remodelers use the energy of ATP to slide or eject nucleosomes, unwrap nucleosomal DNA or exchange dimers to incorporate histone variants (Becker and Horz, 2002; Clapier and Cairns, 2009; Mueller-Planitz et al., 2013). Remodelling enzymes are grouped into four distinct classes (SWI/SNF, ISWI, CHD and INO80) and typically function in the context of large multi-subunit complex where accessory factors control the activity of the catalytic subunit. For example, ISWI that is incorporated in the ACF or CHRAC complexes generates regular nucleosomal patterns that contribute to chromatin assembly and may confer transcriptional repression. In contrast, the ISWI-containing NURF complex randomises nucleosome spacing and promotes transcriptional activation (Deuring et al., 2000; Badenhorst et al., 2002; Fyodorov et al., 2004; Maier et al., 2008; Clapier and Cairns, 2009). Characteristic for all remodelling complexes are protein domains that recognise modified histones, such as bromo- or chromodomains (see 2.3), thereby providing additional layers of regulation by post-translational histone modifications.

In addition to ATP-dependent chromatin remodelers, other mechanisms also regulate chromatin structure directly or indirectly. These principles include the incorporation of histone variants, intrinsic properties of specific DNA sequences, the process of transcription (reviewed in (Segal and Widom, 2009; Korber, 2012)), and post-translational modifications on histones.

2.3 Post-translational histone modifications

Histones are subject to a variety of chemical, post-translational modifications (PTMs). At least 20 different types of chemical modifications have been reported on at least 70 different amino acid residues along the canonical core histones (Kouzarides, 2007; Sakabe et al., 2010; Tan et al., 2011; Xie et al., 2012; Xu et al., 2014). While most modifications are directed towards lysines (acetylation, methylation, ubiquitination, SUMOylation, neddylation, formylation, propionylation, butyrylation, crotonylation, succinylation, malonylation) other amino acids are also modified (arginine methylation; ADP-ribosylation on glutamate and arginine; phosphorylation on serine, threonine and tyrosine; *O*-GlcNAcylation on serine and threonine; proline isomerization and conversion of arginine to citrulline). It is important to note, however, that only a few modification types have been well characterised yet, including lysine methylation and acetylation and serine phosphorylation (Figure 3), and additional work is needed to ascertain the biological relevance of the recently identified marks. Nevertheless, the diversity of modifications and the high number of putative modifiable residues generates an enormous combinatorial potential, which provides the basis for cross-talk among the modifications.

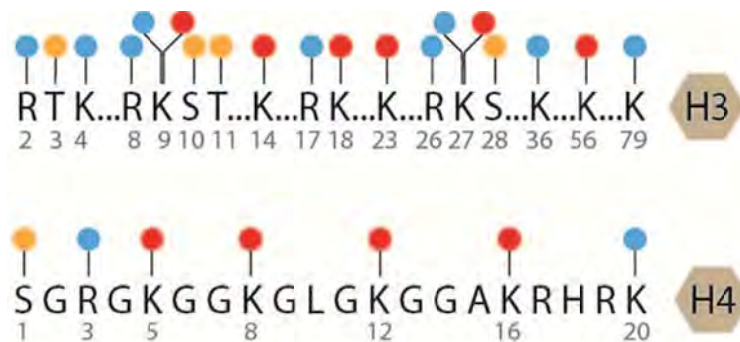


Figure 3. Post-translational modifications on histones H3 and H4.

Major methylation (blue), acetylation (red) and phosphorylation (yellow) sites are shown for histones H3 and H4. Most modifications occur on the N-terminal tail regions. Two intensively studied histone core modifications are acetylation on H3.K56 and methylation on H3.K79.

Depending on the precise modification and residue, three principal modes of action are discussed. First, a modification may have a direct structural impact by modulating DNA-histone or nucleosome-nucleosome interactions. Second, specific modifications or combinations thereof may constitute marks that are bound by proteins with structural or enzymatic activities. Third, a modification may prevent or promote the placement of another mark.

2.3.1 Lysine acetylation

Early studies linking histone acetylation to gene activity

In the early 1960s, seminal work by Vincent Allfrey and others demonstrated that the basic, arginine-rich histone proteins inhibit the transcription of DNA (Allfrey and Mirsky, 1962; Huang and Bonner, 1962; Allfrey et al., 1963). In 1963, lysine acetylation was the first chemical modification described on histones (Phillips, 1963). One year later, after presenting evidence that lysine acetylation and methylation occurs *after translation*, Allfrey surmised: ‘*Such modifications of histone structure, acetylation in particular, may affect the capacity of the histones to inhibit ribonucleic acid synthesis in vivo*’ (Allfrey et al., 1964). In the 1970s, acetylated histones were found at actively transcribed chromatin (Sealy and Chalkley, 1978; Vidali et al., 1978). In the late 1980s, two landmark studies contributed by the labs of Roger Kornberg and Michael Grunstein demonstrated that nucleosomes repress transcription *in vitro* but can relieve repression when ablated *in vivo* (Lorch et al., 1987; Han and Grunstein, 1988). Acetylation was mapped to lysines on the histone termini, thereby providing the rationale that charge neutralization of histones decreases its affinity towards DNA and destabilises nucleosome-nucleosome interactions (Nelson, 1982; Hong et al., 1993). Abrogating acetylation by mutating individual lysine residues on H4 in yeast cells diminished the capacity to induce transcription (Durrin et al., 1991) while the presence of acetylation promoted the binding of transcription activators (Lee et al., 1993; Vettese-Dadey et al., 1996).

Comparing individual and combined lysine site mutations, Grunstein and co-workers derived two conclusions that were strongly influential to the field of chromatin biology and still dominates today’s reasoning of how histone acetylation regulates transcription. In these experiments performed in *S. cerevisiae*, they substituted lysine (K) residues to arginines (R), which mimics the positive charges but cannot be acetylated (Durrin et al., 1991). First, they discovered that the yeast GAL1 promoter is not sufficiently induced in mutants harbouring triple-R or tetra-R mutations whereas single K→R substitutions of the four N-terminal lysine residues (K5, K8, K12, K16) do not diminish GAL1 induction. Based on these observations, they concluded that individual lysine acetylation sites function redundantly by neutralizing the charge state on the histone H4 tail. Following this reasoning, one would predict that mutating all four lysines to glutamines (Q), which resembles lysine structurally and mimics the charge neutralization conferred by acetylation, results in increased promoter activation. However, Grunstein and co-workers observed a tenfold reduction in the GAL1 activity for tetra-Q

mutants, similar to the effects observed in triple-R mutants, suggesting that permanent charge neutralization alone cannot sufficiently explain the observed effects of histone acetylation in transcriptional activation. Second, when they compared how the mutants affect different promoters, they observed that the histone H4 tail is required to activate inducible gene promoters (GAL1, PHO5), but is less critical for rapid response genes (CUP1) and shows no effect on the expression of housekeeping genes (GAL4, PRC1). They concluded that the function of histone acetylation is to induce but not maintain transcription. Further, they speculated that specific gene promoters, such as CUP1, may be depleted of nucleosomes and hence require chromatin-independent induction systems (Durrin et al., 1991).

Regulating transcription by histone acetylation

Different mechanisms are discussed how histone acetylation regulates transcription. As described in the next section, lysine acetylation can directly influence the folding of chromatin at different levels. This includes compromising the stability of individual nucleosomes or interfering with the formation of nucleosomal fibres. Ultimately, the resulting chromatin structure is more permissive, which allows access of transcription factors to the underlying DNA.

In addition to dampening the repressive effect of chromatin directly, histone acetylation can promote transcription by recruiting effector proteins that recognise acetylated lysines via bromodomains (Dhalluin et al., 1999; Filippakopoulos and Knapp, 2012; Sanchez et al., 2014). These domains are frequently found in many co-activators, chromatin remodelers and acetyltransferases and their integrity correlates with the ability to stimulate transcription. In contrast to other histone PTM binding motifs, such as the ones that recognise methylated lysines (chromodomain, PHD finger, tandem tudor domain), bromodomains often display only low affinity towards individually acetylated lysines (Ruthenburg et al., 2007). However, there is emerging evidence that the affinity and specificity of bromodomains is modulated by adjacent modifications (Filippakopoulos et al., 2012). Moreover, multiple acetylation sites may increase the affinity. In a thought-provoking study it was reported that the bromodomain 1 (BD1) of Brdt shows a high affinity towards the di-acetylated mark H4.K5acK8ac (K_D of 22 μ M), while it displayed 10 times lower affinities towards other di-acetyl combinations on the same H4 peptide (Moriniere et al., 2009). This tantalizing result may suggest that certain bromodomains read combinatorial patterns of acetylation motifs rather than individual acetylation marks. This reminds on a structural study that proposes a ‘helical wheel’ organization of the H4 tail (Baneres et al., 1994). According to this model, 100° phasing between individual amino acids of the H4 tail results in a positioning of the four lysines K5, K8, K12 and K16 directly next to each other. Such configuration may provide an excellent configuration to present modifications for combinatorial read-out by bromodomains.

Many examples are documented where histone acetylation synergistically functions with other chromatin modifying principles. For example, while histone acetylation recruits chromatin remodelers that in turn create permissive chromatin structures, additional co-activators bind to newly accessible acetylated lysines and further promote and maintain active transcription (Gregory et al., 1998; Reinke et al., 2001; Kasten et al., 2004; Devaiah and Singer, 2013). Moreover, some individual genes may require a rapid turnover of acetylation instead of a permanently fixed acetylation state for their activation (Hazzalin and Mahadevan, 2005). Lastly, histone acetylation has been shown to mark individual genes for rapid re-activation after mitosis ('gene bookmarking', (Zhao et al., 2011)).

Regulating higher order chromatin structure by histone acetylation

Acetylation of histone H4 provides the best-understood case for how post-translational histone modifications directly regulate higher-order chromatin structure. Central to this function is the interaction between an acidic patch at the H2A-H2B dimer with the basic patch at the H4 tail (amino acids 16-20) (Luger et al., 1997; Zhou et al., 2007). A special role for the H4 tail compared to the other histone tails was already observed early on and was contributed to bridging of adjacent nucleosomes, binding of linker DNA and association with H2B within the same nucleosome (Fletcher and Hansen, 1995; Tse and Hansen, 1997; Hansen et al., 1998; Chodaparambil et al., 2007; Kan et al., 2009; Allahverdi et al., 2011). In a landmark study, Peterson and colleagues used a native chemical ligation strategy to generate nucleosomal arrays with homogenously labelled acetylated lysine 16 on histone H4 (H4.K16ac). These arrays inhibited the formation of inter-fibre contacts (the proposed 30 nm fibre) and reduced nucleosome oligomerization through cross-fibre interactions (Shogren-Knaak et al., 2006). The concept that acetylation of H4.K16 reduces the propensity to form higher-order chromatin structures was reinforced by studies from the Rhodes and Nordenskiöld labs. Using very long nucleosomal arrays (containing 61 copies of the '601' nucleosome positioning sequence instead of 12 copies used by Peterson and colleagues), stoichiometric concentrations of linker histone H5 and only 30% of acetylated H4.K16, the study by Rhodes and co-workers emphasised the dominant role of H4.K16ac over individual or combined H3 and H4 tail deletions (Robinson et al., 2008). Nordenskiöld and colleagues extended these observations by documenting that shorter nucleosomal arrays (12 copies of 601 sequence) with mono-acetylated H4.K16ac have a stronger unfolding effect than either mutating or acetylating the adjacent lysines on H4 (K5, K8, K12) (Allahverdi et al., 2011).

In comparison to acetylation on the H4 tail, our knowledge of how other histone PTMs affect higher-order chromatin structures is rudimentary. Muir and colleagues reported that nucleosomal arrays modified with ubiquitylated lysine 120 on H2B also decrease chromatin compaction in a distinct but synergistic manner to H4.K16ac (Fierz et al., 2011). An inverse effect was reported for nucleosomal arrays decorated with tri-methylated H4.K20, which increased folding and required less divalent magnesium ions to form condensed chromatin (Lu et al., 2008). More subtle effects were observed for

nucleosomes with di-methylated H3.K79, which displayed altered nucleosomal surfaces and consequently may affect binding of effector proteins (Lu et al., 2008). Despite the importance of comparing the effect of different histone PTMs on chromatin folding, it should be stressed that additional parameters including the type and concentrations of salts and highly charged molecules such as polyamines greatly influence the extent of the observed effects and the susceptibility towards individual chemical modifications (Lu et al., 2008; Allahverdi et al., 2011; Liu et al., 2011; Korolev et al., 2012).

In addition to directly affecting the folding of chromatin fibres, histone acetylation modulates the targeting and enzymatic activity of chromatin remodelling complexes. For example, it was proposed that binding of the remodeler subunit BPTF to H4.K16ac and H3.K4me3 through its bromodomain and PHD finger contributes to its localization at specific chromatin regions (Ruthenburg et al., 2011). Acetylation of H3.K56 promotes a switch in the enzymatic activity of the remodeler complex SWR-C to not only incorporate but also remove the histone variant H2A.Z (Watanabe et al., 2013).

An illustrative and complicated case is the relationship between ISWI-containing complexes and acetylation at H4.K16. This histone mark is necessary and sufficient for the strong decondensation phenotype of the male X chromosome in *Drosophila* mutants devoid of ISWI ((Deuring et al., 2000; Corona et al., 2002), see also 2.4 for a detailed discussion on dosage compensation, the underlying process that regulates this male X-chromosome specific phenomenon). However, a series of studies aiming to address the underlying mechanism of this clear phenotypic observation show partly controversial results. While early studies demonstrated that H4.K16ac-containing peptides and mono-nucleosomes moderately inhibited the enzymatic activity of ISWI and the ISWI-containing ACF complex (Clapier et al., 2001; Corona et al., 2002; Shogren-Knaak et al., 2006), this effect was not observed in a recent study using more physiological nucleosomal arrays containing H4.K16ac ((Klinker et al., 2014), and see also (Nightingale et al., 2007)). The situation might be even more complex because ISWI is also required for the incorporation of the linker histone H1 (Lusser et al., 2005; Maier et al., 2008), which in turn counteracts H4.K16ac mediated decondensation (Corona et al., 2007). Moreover, the N-terminus of ISWI resembles the basic patch of H4 and confers autoinhibitory functions, which is relieved by binding to the unmodified H4 tail (Clapier and Cairns, 2012). In summary, these studies illustrate the complications of the *in vitro* approach in dissecting how individual histone modifications influence enzymatic activities on chromatin. *In vivo*, these modifications do not occur in isolation but rather co-occur with other marks along the same histone molecule.

Lysine acetyltransferases

Lysine acetyltransferases transfer an acetyl moiety from acetyl-CoA to the ϵ -amino group of a target lysine. Since their first isolation in the mid 1990s (Kleff et al., 1995; Brownell et al., 1996), dozens of lysine acetyltransferases have been identified and many characterised (Sterner and Berger, 2000; Yang and Seto, 2007; Aka et al., 2011). Because their activities had initially been attributed to histones only, they had been named histone acetyltransferases (HATs). However, with the identification of an increasing number of non-histone substrates, their nomenclature has been updated to lysine acetyltransferases (KATs) (Allis et al., 2007). It is a general notion that KATs are rather promiscuous enzymes that target multiple lysines on histone and non-histone substrates. Many KATs function within the context of multi-subunit protein complexes where accessory subunits confer genomic targeting and influence the substrate specificity (Lee and Workman, 2007).

Based on their sequence homology, KATs are grouped into two main families (GNAT and MYST). In addition, a third class contains proteins with acetyltransferase activity that do not share sequence similarities with the two main families. Prominent members of this third class are the closely related and metazoan-specific co-activators CBP and p300 and the fungal-specific Rtt109.

Despite their sequence divergence, all KATs contain a structurally conserved core region of a three-stranded β -sheet and a long parallel α -helix (Yuan and Marmorstein, 2013). Latter contains four conserved motifs (A-D) that are important for acetyl-CoA binding. The conserved core region is flanked by family-specific arrangement of α -helices and β -sheets. Interestingly, although CPB/p300 and Rtt109 do not show conservation at the sequence level, their structures show a high overall similarity (Wang et al., 2008a). Together, the conserved core region and the flanking segments form a cleft to accommodate the histone target. Both GNAT- and MYST-type KATs contain a conserved glutamate deep within the active centre that acts as a general base for catalysing the nucleophilic attack of the primary target lysine on the thioester bond of acetyl-CoA. In GNAT-type KATs, acetyl-CoA and the histone target first form a ternary complex with the enzyme before catalysis can occur (bi-bi mechanism). MYST-type KATs form first an acetylated intermediate on a conserved cysteine before a glutamate residue facilitates the transfer of the acetyl group from the cysteine to the histone lysine target (ping-pong catalytic mechanism) (Yuan and Marmorstein, 2013). In contrast, p300 does not employ a general base but rather uses a conserved tyrosine as a general acid for catalysis. According to the proposed hit-and-run mechanism, the histone peptide interacts only weakly with the surface of p300, allowing the target lysine to enter into the pocket of the active centre. Together with its less apolar catalytic groove, this may explain the low degree of substrate selectivity for CBP/p300 (Liu et al., 2008).

The Gcn5-related N-acetyltransferase (GNAT) family is the largest group of KATs and includes the well-characterised enzymes HAT1 (KAT1), GCN5 (KAT2A), PCAF (KAT2B) as well as ELP3,

ATAC2, CDY, ECO1 (ESCO1/2) and MEC17. **HAT1** acetylates lysines 5 and 12 on histone H4, which is considered to promote the incorporation of the histone (H3-H4)₂ dimer into chaperone complexes and their subsequent nuclear import. For a detailed discussion on HAT1 functions and substrate specificity, please see 3.4.4. Yeast **GCN5** has long been known as a potent transcriptional co-activator. GCN5 is the catalytic subunit of at least four yeast complexes (SAGA, SLIK, ADA and HAT-A2), two *Drosophila* complexes (SAGA and ATAC) and its two human paralogues GCN5L and **PCAF** incorporate into at least four mutually exclusive complexes (PCAF, STAGA, TFTC and ATAC) (Nagy and Tora, 2007; Spedale et al., 2012). While recombinant yGCN5 acetylates lysines 8 and 16 on H4 and lysine 14 on H3, incorporation into multi-subunit complexes changes its substrate specificities (ADA: H3 lysines K18 and K23; SAGA: H3 lysines 9, 14 and 18) (Grant et al., 1999). The analysis of yeast *gcn5* mutants confirmed the H3 target sites, albeit the reductions varied strongly among the individual sites and showed also decreased H3.K27ac levels (Durant and Pugh, 2006). In contrast, *Drosophila* mutants lacking GCN5 display reduced acetylation levels on lysines 9 and 14 on H3 and 5 and 12 on H4. However, a recent study reported that disruption of the *Drosophila* SAGA complex only reduced H3.K9ac while leaving the other acetylation sites unchanged (Mohan et al., 2014). Interestingly, deleting murine PCAF and GCN5L causes only reduced H3.K9ac levels (Jin et al., 2011), which is in contrast to other studies that reported a number of other putative substrate sites in mammalian cells including H3.K14ac, H3.K18ac and H4 sites (Yang et al., 1996; Herrera et al., 1997; Ogryzko et al., 1998; Zheng et al., 2013b).

The MYST group comprises its founding members human MORF, yeast YBF2 (SAS3), yeast SAS2 and mammalian TIP60 (Esa1 in yeast) as well as MOF, HBO1 and MOZ. **MOF** (males-absent on the first) is an exceptionally specific KAT that presumably targets only a single lysine residue, H4.K16. MOF's role in transcriptional regulation is best understood in *Drosophila* where it is a key subunit of the MSL-DCC complex and essential for X chromosome dosage compensation in male flies. *Drosophila* dosage compensation provides a paradigm for transcriptional regulation by a specific histone acetylation site and will be discussed in detail in section 2.4. During my PhD studies, I contributed to the understanding of MOF's function in a second multi-protein complex, the non-specific lethal (NSL) complex (discussed in sections 3.1-3.4). In addition to regulating transcription, studies conducted in mammals pointed to MOF's role in DNA damage repair, maintenance of pluripotency, autophagy, and apoptosis (Li et al., 2012; Fullgrabe et al., 2013; Yang et al., 2014). Whether these additional functions are exerted in the context of the human MSL, NSL or novel complexes and whether they require the enzymatic activity of MOF is not well understood.

TIP60 and its yeast homolog Esa1 are the catalytic subunit of the NuA4 complex and have been shown to acetylate histone H4 (K5, K8, K12 and K16), H3 (K14) and H2A (K5) as well as numerous non-histone proteins including p53, c-myc, ATM and AR (Sterner and Berger, 2000; Sapountzi and Cote, 2011). TIP60 functions as a co-activator for a wide range of transcription factors, including

NFkB, p53 and nuclear receptors. In addition to regulating transcription, TIP60 has a central role during DNA damage repair. First, acetylation of the DNA damage response protein ATM stimulates the autophosphorylation and activation of this central kinase, leading to phosphorylation of H2A.X, which is a prominent DNA damage signal (Sun et al., 2005). Second, acetylation of H4 renders chromatin accessible and thereby promotes the association of the DNA repair machinery (Tamburini and Tyler, 2005; Murr et al., 2006). Third, during the late steps of repair, TIP60 acetylates phosphorylated H2A.X, which promotes its exchange with the non-phosphorylated H2A.X (Kusch et al., 2004).

The closely related acetyltransferases **MOZ** and **MORF** have central roles during normal development and fusion proteins with CBP are frequently found in human leukemias. In the mouse, MOZ has been shown to acetylate H3.K9ac and is required for the self-renewal capacity of hematopoietic stem cells and the proper development of the thymus and the cardiovascular system (Voss et al., 2012). MORF is required for the self-renewal capacity of adult neural stem cells in the mouse (Merson et al., 2006; Sheikh et al., 2012) and was shown to acetylate H4 at lysines 5, 8, 12 and 16 (Champagne et al., 1999).

Studies on the acetyltransferase **HBO1** in different experimental setups such as human cell lines and *hbo1* null mouse mutants reported conflicting results with regard to its role during DNA replication, cell cycle regulation and histone substrate selection (see discussion in 3.4.4). In summary, the MYST family is a functionally diverse class of acetyltransferases that are implicated in the regulation of transcription and DNA damage repair and required for proper development.

The two closely related acetyltransferases **CBP** and **p300** (one member in *Drosophila*: dCBP/nejire) are ubiquitously expressed potent transcriptional co-activators that have been shown to acetylate a wide range of histone sites (H2A: K5; H2B: K12, K15; H3: K14, K18, K23, K27, K56; H4: K5, K8, K12, K16) (Bannister and Kouzarides, 1996; Ogiwara et al., 2011; Wang et al., 2013b). In addition, they form a platform to recruit other factors involved in transcriptional regulation, including other acetyltransferases (GCN5, PCAF), sequence-specific DNA-binding transcription factors (among them p53, c-myc, FOXO3a), general transcription factors (TFIIB) and nuclear hormone receptors (ER, AR) – many of these proteins are also potent acetylation substrates for CBP/p300 (Wang et al., 2013b). The potential of CBP/p300 to stimulate transcription seems to be context-dependent. For example, while genes regulated by the nuclear hormone receptor PPAR strictly depend on the HAT activity of CBP/p300, other CBP/p300 bound genes rely on the ‘platform’ function to recruit other activators but do not require the enzymatic activity of these KATs (Jin et al., 2011; Bedford and Brindle, 2012). Not surprisingly, *cbp* or *p300* null mice show early embryonic lethality and their deficiencies or deregulation in humans contributes to cancer, neurodegenerative disorders and heart disease (Bedford and Brindle, 2012; Valor et al., 2013).

Lysine deacetylases

Lysine acetylation is reversed by lysine deacetylases (KDACs). These enzymes are grouped based on sequence homology, phylogenetic analysis and co-factor dependency in four classes. The zinc-dependent ‘classical’ deacetylases, biochemically isolated first in 1996 (Taunton et al., 1996) and named according to their initially described substrates as histone deacetylases (HDACs) are sorted in class I (HDACs 1-3 and 8), class II (HDACs 4-7, 9 and 10) and class IV (HDAC 11). The NAD⁺-dependent deacetylases of class II are named according to their founding member silent-information regulator-2 (Sir2, (Imai et al., 2000; Landry et al., 2000; Smith et al., 2000)) as Sirtuins and contain seven members in humans (SIRT1-7). In comparison to lysine acetyltransferases, deacetylases display an even more relaxed substrate specificity directed against many acetyl-lysines in histones and non-histone proteins (Feldman et al., 2012; Rauh et al., 2013).

The class I HDACs function as classical transcription repressors that promote repressive chromatin structures by deacetylating all four core histones (Yang and Seto, 2008). The four mammalian class I HDACs (two in *Drosophila*: HDAC1 and 3) are homologs of the yeast transcription repressor RPD3. HDAC1 and HDAC2 are highly similar proteins, interact with each other and form the catalytic core of three well-defined multi-protein complexes (SIN3, NURD and CoREST). NURD and CoREST contain SANT-domain containing proteins, which bind histones and stimulate the histone deacetylase activity (Cunliffe, 2008). In addition, other chromatin-modifying activities and histone PTM-recognizing proteins are integral components in these deacetylase complexes, providing the basis for coupling histone deacetylation with histone lysine demethylation (LSD1 within CoREST), nucleosome remodelling (Mi-2 within NURD) and recognition of methylated DNA (MBD2 within NURD) and histones (mammalian ING2 and yeast EAF3 within SIN3). HDAC3 also shows high sequence similarity to HDAC1/2 but contains only a single instead of two 14-3-3 phosphorylation site binding motifs. HDAC3 is the catalytic subunit of the N-CoR/SMRT co-repressor complexes that in addition also contain the lysine demethylase JMJD2A/KDM4A. In contrast to the other class I HDAC members, HDAC8 has not yet been found in a stable protein complex and its functions are less well characterised. Recent studies indicated that HDAC8 interacts with multiple components of the cohesion complex (Joshi et al., 2013) and it shows deacetylase activity against one cohesion subunit (SMC3, (Deardorff et al., 2012)).

Mammalian class II HDACs are related to yeast Hda1 and are grouped into class IIa (HDACs 4, 5, 7 and 9) and class IIb (HDACs 6 and 10). Class IIa HDACs contain multiple conserved phosphorylation motifs at their N-termini and a central deacetylase domain. These phosphorylation motifs are critical for the nucleocytoplasmic shuttling of these enzymes and their function as essential signal transducers downstream of many cytosolic signalling cascades. Vertebrate class IIa HDACs show a very low intrinsic deacetylase activity that is caused by a substitution of tyrosine to histidine within the catalytic domain (Lahm et al., 2007). In line, the catalytic activity purified from cellular HDAC4 and HDAC7

likely results from its association with HDAC3 (Fischle et al., 2001; Fischle et al., 2002). Interestingly, the single class IIa HDAC in *D. melanogaster*, HDAC4, does not contain this substitution, suggesting that the enzyme has a robust deacetylase activity in the fly. HDAC6 contains a tandem deacetylase domain and a C-terminal zinc finger, predominantly localises to the cytosol, where it deacetylates numerous proteins including α -tubulin, cortactin and HSP90. In comparison to the other zinc-dependent deacetylases, the functions of the vertebrate-specific class IIb HDAC10 and the highly conserved class IV HDAC11 are less well understood.

Members of the sirtuin family of class III deacetylases couple lysine deacetylation with NAD^+ hydrolysis to yield nicotinamide and O-acetyl-ADP-ribose in addition to the non-acetylated product. The requirement for NAD^+ directly links the enzymatic activity of sirtuins to the energy status of the cell. In line, sirtuins have received much attention as critical regulators for metabolic processes including glycolysis, gluconeogenesis and fat oxidation and higher levels of SIRT1 were initially reported to mediate the effects of caloric restriction that increases life span in yeast, *C. elegans* and *D. melanogaster*. More recent studies, however, questioned the capacity of lifespan extension of SIRT1 in flies (*sir2*) and worms (*sir-2.1*) (reviewed in (Feldman et al., 2012; Sauve and Youn, 2012)). Of note, while SIRT1 knockout mice show normal lifespan, SIRT6 deficient mouse mutants display phenotypes of accelerated aging. Sirtuins show distinct localization patterns in mammalian cells, where SIRT1, SIRT6 and SIRT7 are pre-dominantly found in the nucleus, SIRT2 in the cytosol, and SIRT3-5 in the mitochondria (Verdin et al., 2010). In addition to their deacetylase activity, individual members have been suggested to remove other acyl groups including malonyl and succinyl (Du et al., 2011). Moreover, SIRT4 and SIRT6 display ADP-ribosyltransferase activity. Indeed, until recently, no deacetylation reaction had been reported for SIRT4, yet a recent acetylome microarray study suggested multiple deacetylation targets for this enzyme (Rauh et al., 2013).

A reoccurring theme is that KATs and KDACs regulate the enzymatic activity of each other and of other chromatin-modifying activities. For example, SIRT1 mediated deacetylation of SUV39H1 increases the catalytic activity of this methyltransferase and is necessary for proper heterochromatin formation (Vaquero et al., 2007). Likewise, autoacetylated hMOF is deacetylated by SIRT1, which increases the affinity of MOF towards nucleosomes *in vitro* and is necessary for target gene binding *in vivo* (Lu et al., 2011). In turn, hMOF acetylates the SIRT1 inhibitor protein DBC1 to generate more active SIRT1 molecules. In a negative-feedback loop, SIRT1 deacetylates its own inhibitor and thereby restricts its activity (Zheng et al., 2013a).

2.3.2 Lysine methylation

Lysine methylation is catalysed by lysine methyltransferases (KMTs, formerly: histone methyltransferases (HMTs)) and reversed by lysine demethylases (KDMs, formerly: histone demethylases (HDMs)). These enzymes add or remove up to three methyl-groups and show a high specificity towards the specific lysine residue and degree of methylation. All lysine methyltransferases require S-adenosyl methionine (SAM) as a methyl-donor and most KMTs share a specific catalytic domain (SET domain: Su(var)3-9, Enhancer of Zeste, Trithorax). Histone demethylases are classified according their catalytic mechanism in flavin adenine dinucleotide (FAD)-dependent amine oxidase, and Fe(II) and α -ketoglutarate-dependent dioxygenase (Smith and Denu, 2009). Genome-wide localization studies using chromatin-immunoprecipitation combined with microarray or deep-sequencing (ChIP-chip, ChIP-Seq) revealed that lysine methyl marks can be grouped into active (H3.K4me3, H3.K36me3, H3.K79me2) and repressive (H3.K9me3, H3.K27me3, H4.K20me3) signatures. Because methylation does not change the positive charge state of lysines, its primary function is likely to recruit or oppose specific effector proteins. Those effector proteins recognise methylated lysines through highly dedicated protein folds, including chromodomains, PHD fingers and MBT domains and thereby connect lysine marks to a whole range of downstream processes (Taverna et al., 2007).

Tri-methylation of lysine 4 on H3 (H3.K4me3) is a hallmark of gene promoters. This mark may stimulate transcription via the recruitment and/or stabilization of the transcription factor TFIID, the chromatin remodeler BPTF/NURF and the H3.K9me2-demethylase PHF8 (Taverna et al., 2007; Herz et al., 2013). H3.K4me3 is catalysed by a single enzyme in yeast (SET1) but three enzymes in *Drosophila* (SET1, TRX, TRR) and at least six enzymes in mammalian cells (MLL1-4, SETD1A, SETD1B). The expansion of this family in higher eukaryotes indicates a high level of specification and potentially redundancy. In line with this, MLL1 pre-dominantly localises to the *hox* gene cluster and *mlL1* knockout mice display characteristic homeotic phenotypes (Yu et al., 1995). Compared with the highly specific lysine methyltransferases, the catalytic activity of many lysine demethylases is strongly influenced by their interacting proteins. Consequently, individual demethylases do not only revert specific methylation states on H3.K4 but also act on other lysine residues (for example KDM2B: H3.K4me3 and H3.K36me1/2, LSD1: H3.K4me2/3 and H3.K9me1/2) (Hojfeldt et al., 2013).

In contrast to H3.K4me3, the mono-methylated form (H3.K4me1) is not restricted to promoters but it is broadly distributed across the genome and enriched on the bodies of active genes and enhancers (Herz et al., 2013). In a recent study, Shilatifard and co-workers could demonstrate that *Drosophila* mutants lacking the responsible enzyme for H3.K4me1 (TRR) also lose the enhancer-mediated gene activation of the *cut* locus. Moreover, they observed a global decrease of the enhancer mark H3.K27ac. Together with previous studies that showed an interaction between the H3.K27me3-

demethylase UTX (KDM6A) with TRR (MLL3/MLL4), these results support the model that an acetylation-methylation switch at H3.K27 regulates the transition from an inactive but poised to an active enhancer (Herz et al., 2012).

Methylation of H3.K36 and H3.K79 is found on actively transcribed gene bodies. Similar to the enzymes that catalyse H3.K4 methylation, there is only a single yeast enzyme (SET2) that sets all three methylation states of H3.K36 whereas three *Drosophila* (SET2, MES-4 and ASH1) and at least six mammalian (SET2, ASH1L, SETMAR, NSD1-3) enzymes preferentially generate either the mono- and di-methylated states or the tri-methylated state, respectively. Early studies suggested that H3.K36me3 is recognised by the chromodomain of EAF3 to recruit the RPD3S complex, which in turn deacetylates histones at gene bodies and thereby suppresses cryptic transcription (Carrozza et al., 2005; Keogh et al., 2005). However, a recent report demonstrated that RPD3S still localises to genes in the absence of H3.K36me3 – yet it is inactive – suggesting that this mark promotes the activity rather than the recruitment of a deacetylase complex (Drouin et al., 2010).

Methylation to H3.K79 is unique in many instances. First, it is catalysed by a non-processive methyltransferase (DOT1L) that does not contain a SET domain but structurally is related to arginine methyltransferases (Min et al., 2003). Second, although there is evidence for dynamic changes of H3.K79me2 during the cell cycle, no demethylase is known that actively removes this mark (Nguyen and Zhang, 2011). Third, despite being present at all active gene bodies, only a few genes seem to require this mark for their activity. Instead, accumulating evidences point towards a role in DNA repair (via BP53 recruitment) and activation of the G1/S checkpoint (Huyen et al., 2004; Wysocki et al., 2005; Nguyen and Zhang, 2011).

Di- and tri-methylated H3.K9 is a classical marker for constitutive heterochromatin found at genomic regions that contain repetitive DNA elements. In a self-reinforcing spreading mechanism, SUV39H1/2 catalyse H3.K9me2/3 that is subsequently bound by HP1- α , which in turn binds and stabilises SUV39H1/2 at these regions. In addition, HP1 also recruits SUV420H1, which catalyses H4.K20me3 to confer further condensation. Cells lacking the SUV39 and SUV420 enzymes display derepressed satellite transcription, chromosome translocation and mitotic defects (Peng and Karpen, 2009; Black et al., 2012; Jorgensen et al., 2013). In addition to repeat DNA silencing and genomic integrity, H3.K9me2 also represses genes that are covered by large megabase-scale domains found in differentiated but not pluripotent cells (Wen et al., 2009). Similar to the HP1- α -SUV39H1/2 feedback loop at pericentric heterochromatin, these domains may be established and stabilised by a self-reinforcing process that involves di-methylation of H3.K9 by G9a followed by stabilization of the enzyme via binding of G9a to H3.K9me2 (Shinkai and Tachibana, 2011). Interestingly, in contrast to most other chromatin enzymes, H3.K9 methylating enzymes do not form stable multi-subunit assemblies but rather engage in transient interactions. For example, while the methyl-CpG binding protein MBD1 recruits the H3.K9 methyltransferase SETDB1 into a S-phase specific complex with

CAF-1 (Sarraf and Stancheva, 2004), the co-repressor KAP-1 directs SETDB1 to silence individual genes (Schultz et al., 2002).

Methylation to H3.K27 is generally associated with gene repression. H3.K27me1 is found at constitutive heterochromatin together with H3.K9me3 yet its function is not well understood. In addition, low levels of H3.K27me1 are detected at active genes that presumably arise from the demethylation of H3.K27me2/3 by UTX or JMJD3 (Swigut and Wysocka, 2007). H3.K27me2/3 are prominent marks of facultative heterochromatin that are set by the polycomb repressive complex PRC2. ChIP-chip/Seq experiments documented several binding modes of these histone modifications and their respective enzymes (EZH1/2). Similar to H3.K9me2/3, these marks span extended domains of several 100 kb in differentiated but not pluripotent cells. In addition, individual genes (in particular developmental genes), intergenic regions (potentially coding for ncRNAs), subtelomeric regions and LTR retrotransposons are targeted by these modifications (Hawkins et al., 2010). Mechanistically, H3.K27me3 represses transcription by antagonizing the acetylation mark on the same residue and by providing a platform for the PC subunit of the PRC1 complex. Transcription repression by the PRC1 complex is achieved by compacting nucleosomal arrays (via PSC in *Drosophila* and CBX in vertebrates) and ubiquitylation of H2A.K119 by the PRC1 subunit RING1. However, how H2A.K119ubi impedes transcription is not yet known. In addition, there is accumulating evidence that PRC1/2 interferes with the transcriptional machinery. However, whether this occurs on the level of RNA polymerase II recruitment, promoter-proximal release or by inhibiting mediator targeting is controversially discussed (Chopra et al., 2011; Margueron and Reinberg, 2011; Lehmann et al., 2012; Simon and Kingston, 2013).

2.4 Dosage compensation in the fruit fly *Drosophila melanogaster*

Dosage compensation describes the process of adjusting transcription levels between the unequal numbers of sex chromosomes in heterogametic species. Although different species have evolved different strategies to balance gene expression between the sexes, they all utilise chromatin-based mechanisms. In mammalian females, one of the two X chromosomes is inactivated by a concerted step-wise process including recruitment of the polycomb silencing machinery via the long non-coding RNA Xist (X-inactive specific transcript) followed by the incorporation of the histone variant macroH2A, deacetylation of histones and DNA hyper-methylation (Wutz, 2011; Jeon et al., 2012; Dupont and Gribnau, 2013). In the nematode worm *Caenorhabditis elegans*, the dosage compensation complex contains components of the meiotic/mitotic condensin complex, which halves the transcription output of the two hermaphrodite X chromosomes (Meyer, 2005). *Drosophila* also employ a subtle but vital transcriptional adjustment, but there, dosage compensation occurs in males to double the transcription from the single X chromosome (Straub and Becker, 2007). Male flies have evolved a unique ribonucleoprotein complex that contains the activities from an acetyltransferase, an E3 ligase and a helicase (see below). Recent studies suggest that dosage compensation in mammals and nematodes involves a second mechanism, that, similar to male flies, upregulates the expression from the X chromosome(s) in order to balance the expression towards the autosomes (Deng et al., 2014). The molecular mechanisms of this newly identified principle are largely elusive and are just beginning to be unveiled (Deng et al., 2013).

Research on dosage compensation in *Drosophila* has a long history. The geneticist Hermann Joseph Muller introduced the term ‘dosage compensation’ based on observations made in the 1920s and 1930s, where the single gene copy of an eye colour marker encoded on the male’s X chromosome produces almost the same eye colour as the genes from the two female X chromosomes (Bridges 1922, Muller 1931, Muller 1950). In 1964, the cytologist Rudkin showed that the single X chromosome in males has a comparable ‘thickness’ to the paired female X chromosome, suggesting enhanced transcriptional activity in males (Rudkin 1964). One year later, Mukherjee and Beerman confirmed this prediction. Incubating salivary glands with tritiated uridine allowed them to observe a similar incorporation between the single male and the two female X chromosomes (Mukherjee and Beermann, 1965). Detecting similar levels of activity from the X-linked gene glucose-6-phosphate dehydrogenase between male and male flies lend further support (Komma, 1966). Together, these experiments establish that a transcription-based process calibrates the gene expression output between the single male and the two female X chromosomes.

Elegant genetic experiments identified the components of the molecular machinery that mediates dosage compensation. A series of genetic screens identified four autosomal loci that, when mutated, killed male but not female flies and reduced the rate of X chromosome transcription (Fukunaga et al., 1975; Belote and Lucchesi, 1980b, a). The genes were named according to their **male-specific lethal**

phenotype, *msl1*, *msl2*, *msl3* and *mle* (*maleless*). Immunological stainings of polytene chromosomes with sera raised against the MSL proteins demonstrated highly overlapping and almost exclusive localization of those factors to the X chromosome, corroborating their function to act directly on this chromosome. The observation that an acetylated histone isoform, H4.K16ac, also predominantly localised to the male X chromosome not only directly suggested a potential mechanism for how dosage compensation works but also provided an intuitive correlation and beautiful visualization of increased acetylation and transcriptional hyper-activation (Turner et al., 1992). Soon thereafter, the enzyme that catalysis this histone mark was identified in a genetic screen for male-lethality causing mutations on the X chromosome: *males absent on the first* (*mof*) encodes a gene product with a recognizable MYST-type acetyltransferase domain (Hilfiker et al., 1997; Gu et al., 1998) and specificity towards acetylating H4.K16 (Akhtar and Becker, 2000). Around the same time, two non-coding RNAs (RNA on the X, *roX1* and *roX2*) were found to be exclusively expressed in male cells, where they are only detectable at the male X chromosome (Amrein and Axel, 1997; Meller et al., 1997). The ribonucleoprotein complex MSL-DCC, containing the five MSL proteins and two *roX* RNAs, forms only in male flies, because *msl2* expression is inhibited in female cells by the action of the master sex regulator *sex-lethal* (Kelley et al., 1995; Bashaw and Baker, 1997; Kelley et al., 1997).

The MSL-DCC binds X chromosomal active genes in a multi-step process. According to the prevalent model, a core complex containing MSL1 and MSL2 binds around 250 ‘high-affinity’ sites (HAS) located in coding regions evenly distributed across the 22 mega base pairs of the X chromosome (Gelbart and Kuroda, 2009; Conrad and Akhtar, 2011; Straub and Becker, 2011). MSL1 acts as a scaffold to allow binding of MOF and MSL3. This results in a full MSL-DCC complex that is capable to spread in cis from the high affinity sites to active genes. Several factors may stimulate the spreading process. First, MOF mediated acetylation of H4.K16 is required for efficient spreading (Gu et al., 1998). Second, efficient loading of the *roX* RNAs by the helicase activity of MLE facilitates the transfer of the complex from HAS to active genes (Morra et al., 2011). Third, mutating the enzyme that places H3.K36me3 on gene bodies lead to reduced binding of the MSL-DCC to active genes while its localization on HAS was not disturbed (Larschan et al., 2007; Bell et al., 2008). However, while an initial study suggested that MSL3 binds H3.K36me3 using its own chromodomain (Sural et al., 2008), subsequent structural and biochemical studies could not confirm a direct interaction between them. Rather, they suggested that the chromodomain of MSL3 binds H4.K20me1 (Kim et al., 2010; Moore et al., 2010). How the MSL-DCC recognises and utilises H3.K36me3 for spreading remains elusive. Fourth, the precise levels of the MSL proteins are important. Excess of subunits lead to ectopic binding events at lower affinity sites present on the autosomes, whereas decreased MSL protein levels (or absence of the spreading factors MSL3 and MOF) reduced binding to active genes (Kelley et al., 1999; Dahlsveen et al., 2006; Straub et al., 2008). Remarkably, the MSL-DCC auto-regulates the correct amounts of its subunits. MSL2 was recently shown to possess an E3 ligase

activity that marks excess subunits with polyubiquitylation and thereby promotes subsequent degradation by the proteasome (Villa et al., 2012). Moreover, absence of MSL2 (or MSL1) increases the spatial distance between HAS measured by 3D fluorescence in situ hybridization (3D-FISH) experiments (Grimaud and Becker, 2009). Together, these observations lead to the current model for selective X chromosome targeting and spreading of the MSL-DCC: Transcription from the two *roX* loci provides a seed for the self-assembly of a dosage compensated domain within the X chromosome. Active, dosage compensated genes are located in the interior of this domain, whereas inactive genes that are not bound by the MSL-DCC reside at the periphery. A gradient of MSL-DCC complexes originates in the core of this domain and diffuses along a gradient of high and low affinity sites towards the periphery. Adjusting precise complex levels restricts its targeting exclusively to the X chromosome (Grimaud and Becker, 2010).

Although increased transcription from the male X chromosome was the founding observation that defined ‘dosage compensation’, the precise molecular mechanism that mediates ‘two-fold up’ has remained largely mysterious until today. High-resolution ChIP mapping of the MSL proteins and the H4.K16ac mark showed localised binding to the gene bodies of active genes, with a tendency to enrich towards the gene’s 3’ end (Smith et al., 2001; Alekseyenko et al., 2006; Gilfillan et al., 2006). Based on this observation, Lucchesi and colleagues suggested the model that dosage compensation occurs at the level of transcriptional elongation. Conceivable, local decondensation of chromatin caused by H4.K16ac marked nucleosomes may facilitate the transition of RNA polymerase II through an otherwise repressive chromatin template (Smith et al., 2001). Two recent studies largely supported this model by documenting increased occupancy of RNA polymerase (or nascent transcripts arising from it) over the gene bodies but not at the promoters (Larschan et al., 2011; Ferrari et al., 2013). Still, how increased acetylation of H4.K16ac brings about a fine-tuned transcriptional stimulation in the two-fold range has remained elusive. Interestingly, Akhtar and Becker observed early on that tethering MOF to a yeast promoter boosts transcription far beyond the two-fold range (Akhtar and Becker, 2000), indicating that additional mechanisms are in place that coordinate a fine-tuned response to arrive at the final ‘two-fold up’.

3 RESULTS AND DISCUSSION

3.1 The activation potential of MOF is constrained for dosage compensation

Matthias Prestel, Christian Feller, Tobias Straub, Heike Mitlöhner and Peter B. Becker

3.1.1 Summary, significance and own contribution

Summary and significance

Although it is established that acetylation of H4.K16 by MOF is required for dosage compensation, it is unknown how the precise two-fold stimulation is achieved. At the time when we started this project, the Akhtar group had recently identified a novel MOF-containing complex in male cells, the non-specific lethal (NSL) complex, which contains two components of the nuclear pores and a number of poorly characterised proteins, including NSL1 and MBD-R2 (Mendjan et al., 2006). In their subsequent work, they demonstrated that MOF binds genome-wide to many promoters in male and female cells, suggesting novel roles for MOF independent of dosage compensation (Kind et al., 2008). The objective of this study was to compare the molecular context and effect on transcription of MOF in male and female flies.

Combining affinity-purification mass spectrometry, genome-wide mapping and transcriptome studies and the analysis of reporter loci in transgenic flies and cell systems, we arrived at four main conclusions. First, in female cells, MOF resides in a similar complex as described before by the Akhtar group, containing NSL1, NSL2, NSL3, MBD-R2 and MCRS2, but lacking the nuclear pore proteins. Second, in male and female cells, the NSL complex (using MBD-R2 as a marker) binds at promoters of many active genes along all chromosomes. Ablation of MBD-R2 reduced transcription of target genes genome-wide and at a defined reporter locus, indicating that the NSL complex is a global transcription activator. Third, in male cells, MOF distributes dynamically between the MSL-DCC and NSL complex. Introducing low levels of MSL2 in female cells causes a global redistribution of MOF away from promoters of genes on all chromosomes towards gene bodies on the male X chromosome. Fourth, the strong activation potential of MOF-mediated H4.K16ac is constrained in the context of dosage compensation. Reconstituting the different targeting principles of MOF at defined reporter loci in transgenic flies revealed that MOF within the NSL complex activates transcription strongly in a distance-dependent manner, typical for promoter-type transcriptional activators. In contrast, MOF recruitment in the context of the MSL-DCC promoted a distance-independent stimulation of transcription in the two-fold range, reminiscent for dosage compensation. Importantly, the reporter locus was decorated with similar levels of H4.K16ac in both conditions, suggesting that additional factors constrain the strong activation potential of this histone mark. In line with these observations, we observed a positive correlation between the levels of H4.K16ac and the gene expression strength for genes in both sexes and on all chromosomes, except for the male X chromosome. Remarkably, depriving MSL2 in male tissue-culture cells increased the expression of the MOF-regulated reporter gene. These observations led us to propose the model that the MSL-DCC harnesses a strong activator (MOF through H4.K16ac) and dampens its activity to achieve a precise two-fold stimulation characteristic for dosage compensation.

The molecular entity that constrains the activation potential of MOF/H4.K16ac is unknown but several repressive chromatin components, which have previously been genetically linked to dosage compensation (Deuring et al., 2000; Spierer et al., 2005; Spierer et al., 2008), provide attractive candidates for such activities. In addition, advancing our understanding of the recently identified ‘over compensating males’ (*ocm*) gene (Lim and Kelley, 2013) and characterizing the molecular targets of the E3 ligase MSL2 may provide further insights (see discussion in section 4). In summary, this report defines an entry point to study the contribution of repressive activities to the process of dosage compensation, which promises to refine our understanding of the general principles that regulate genome balancing at the level of chromosome-wide transcription control.

Own contribution

For this study led by Dr. Matthias Prestel, I designed and performed all genomic experiments and contributed to the characterization of the reporter loci by ChIP-qPCR in adult male and female flies. I prepared the main figures 2F, 3, 4, 5 and 7, the supplementary figures S3 to S6 and contributed to the writing of the manuscript.

3.1.2 Published manuscript

The Activation Potential of MOF Is Constrained for Dosage Compensation

Matthias Prestel,¹ Christian Feller,¹ Tobias Straub,¹ Heike Mitlöhner,¹ and Peter B. Becker^{1,*}

¹Adolf-Butenandt-Institute and Centre for Integrated Protein Science, Ludwig-Maximilians-University, 80336 Munich, Germany

*Correspondence: pbecker@med.uni-muenchen.de

DOI 10.1016/j.molcel.2010.05.022

SUMMARY

The H4K16 acetyltransferase MOF plays a crucial role in dosage compensation in *Drosophila* but has additional, global functions. We compared the molecular context and effect of MOF in male and female flies, combining chromosome-wide mapping and transcriptome studies with analyses of defined reporter loci in transgenic flies. MOF distributes dynamically between two complexes, the dosage compensation complex and a complex containing MBD-R2, a global facilitator of transcription. These different targeting principles define the distribution of MOF between the X chromosome and autosomes and at transcription units with 5' or 3' enrichment. The male X chromosome differs from all other chromosomes in that H4K16 acetylation levels do not correlate with transcription output. The reconstitution of this phenomenon at a model locus revealed that the activation potential of MOF is constrained in male cells in the context of the DCC to arrive at the 2-fold activation of transcription characteristic of dosage compensation.

INTRODUCTION

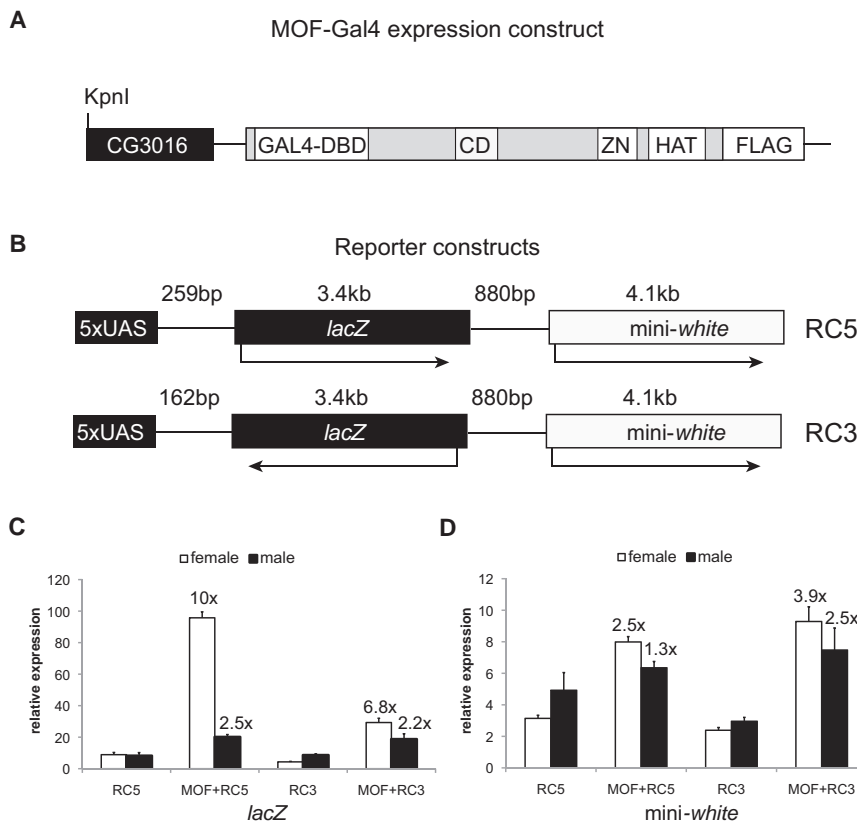
The organization of chromatin affects all aspects of gene transcription. The most prominent chromatin features that can be correlated with particular functional states are various histone modifications. Many of these modifications affect chromatin structure indirectly by providing binding sites for modulators of local chromatin organization (Fischle, 2008). By contrast, the acetylation of histone H4 at lysine 16 (H4K16ac) is known to affect chromatin structure directly by preventing the compaction of the nucleosomal chain into 30 nm fibers (Robinson et al., 2008; Shogren-Knaak et al., 2006). Accordingly, targeting H4K16 acetylation to a promoter in vitro or in yeast can lead to profound derepression of transcription (Akhtar and Becker, 2000).

H4K16ac is the major modification involved in the process of dosage compensation in *Drosophila*. Dosage compensation mechanisms counteract the adverse effects of sex chromosome aneuploidy in a variety of organisms (Lucchesi et al., 2005). In *Drosophila*, like in mammals, females are characterized by two X chromosomes, whereas males only have one X and a gene-

poor, degenerate Y chromosome. The halved dose of X-linked genes in males is counteracted by elevating their transcription levels by roughly 2-fold (Hamada et al., 2005; Straub et al., 2005a). This subtle yet vital adjustment of gene expression depends on the acetyltransferase MOF, which preferentially acetylates H4K16 on the X chromosome in males (Akhtar and Becker, 2000; Hilfiker et al., 1997; Smith et al., 2000). MOF is part of a dosage compensation complex (DCC, also known as male-specific-lethal [MSL] complex), a ribonucleoprotein assembly that specifically associates with the X chromosome. The DCC also contains the MSL proteins MSL1, MSL2, and MSL3; the RNA helicase maleless; and two noncoding RNAs. The DCC only forms in males due to the male-specific expression of MSL2. Targeting to the X chromosome is a multistep process involving the recognition of a relatively low number of high-affinity sites (HAS) or chromosomal entry sites (CES) by MSL1-MSL2 followed by transfer to transcribed gene sequences (reviewed in Gelbart and Kuroda, 2009; Lucchesi et al., 2005; Straub and Becker, 2007). The phenomenon of dosage compensation presents a unique opportunity to study mechanisms of coregulation of genes on a chromosome-wide scale.

Current models for dosage compensation assume that enrichment of H4K16ac on the male X chromosome leads to decompaction of the chromatin fiber, which may facilitate transcription elongation. However, several observations indicate that the role of MOF in dosage compensation is more complex. First, ectopic expression of MOF in yeast leads to a much stronger activation of a reporter gene than the 2-fold effects that characterize dosage compensation (Akhtar and Becker, 2000). The principles that harness the activation potential of MOF on the male X chromosome in *Drosophila* are unknown. Further, MOF is expressed in male and female cells (Hilfiker et al., 1997). Low levels of MOF can be detected at gene-rich interbands of all polytene chromosomes in female larvae and on male autosomes (Bhadra et al., 2000; Kind et al., 2008). MOF was originally isolated due to the male-specific lethality of its loss-of-function phenotype (Hilfiker et al., 1997). However, MOF mutant females are developmentally delayed and have a decreased fertility (Gelbart et al., 2009). Akhtar and colleagues recently suggested roles for MOF in gene expression beyond dosage compensation (Kind et al., 2008). To which extent MOF can catalyze H4K16 acetylation outside of the context of the DCC is controversial (Gelbart et al., 2009; Kind et al., 2008; Morales et al., 2004).

The issue is further complicated by the observation that MOF can be part of an alternative, "NSL" complex that was isolated from male tissue culture (SL2) cells or mixed-sex embryos



(Mendjan et al., 2006). In this context, MOF associates with several poorly characterized factors with interesting domain organization, such as the “nonspecific-lethals” (NSL1, NSL2, and NSL3), dMCRS1, and MBD-R2 (Mendjan et al., 2006).

To gain more insight into the diverse molecular context of MOF function, we affinity purified a MOF complex from female cells that shares many subunits with the “NSL complex” described by the Akhtar group, including MBD-R2. Genome-wide transcriptome studies show that MBD-R2 is a general facilitator of transcription. Combining chromatin immunoprecipitation with probing DNA microarrays (ChIP-chip), we mapped the chromosome-wide interactions of MOF, MSL1 (a marker for the DCC), and MBD-R2 (a marker for the alternative complex) in adult flies sorted according to sex. MOF colocalizes with MBD-R2 at the 5' end of most active genes in females. In males, most of MOF is found in the context of the DCC on the coding regions of active X chromosomal genes. Ectopic reconstitution of the DCC in female cells relocalizes MOF to the X chromosome, which suggests that MOF distributes in a dynamic equilibrium between the two complexes.

Targeting MOF via a heterologous DNA-binding domain to an autosomal reporter locus in flies allowed contrasting the activation potential of MOF in either sex. In the context of MBD-R2 in females, MOF was a potent activator of transcription from nearby promoters. In the context of the DCC, this activation was limited to 2-fold, reminiscent of dosage compensation. Ablation of MBD-R2 and of MSL2 in tissue culture cells allowed

adjusting the relative levels of the two complexes. Interestingly, reduction of MSL2 led to increased transcription. Taken together, our data suggest that dosage compensation involves a hitherto unappreciated principle that constrains the activation potential of this acetylation mark.

RESULTS

Construction of Model Loci to Study the Sex-Specific Regulatory Potential of MOF

To dissect the function of MOF in the context of the intact organism, we employed fly lines bearing reporter gene loci to which regulatory proteins can be recruited (Figure 1). One reporter locus contains binding sites for the yeast transcription factor Gal4 (“upstream activating sequences”; UAS_{Gal}) upstream of *lacZ* and *mini-white* genes (RC5, Figure 1B [Zink and Paro, 1995]). Basal transcription of the *lacZ* gene from a minimal promoter can be boosted by activator binding to the UAS element. The *mini-white* gene, which is transcribed from its own promoter, has been shown to be responsive to dosage compensation in an appropriate X chromosomal context (Qian and Pirrotta, 1995). However, chromosome-wide DCC interaction studies (Straub et al., 2008) showed that the *white* gene does not contain a high-affinity binding site for the DCC (HAS), and the reporter cassette in the autosomal context only shows background binding of MOF in both sexes (see below). A second set of transgenic flies was generated that express MOF fused to

a GAL4 DNA-binding domain (MOF-Gal4; Figure 1A). Crossing reporter and activator lines leads to recruitment of MOF to the reporter locus in the offspring.

We reasoned that the tethering site might substitute for a HAS, which are frequently found outside of coding regions. According to the prevalent model, these sites serve to distribute DCC to close-by transcribed genes (Alekseyenko et al., 2008; Straub et al., 2008). The well-defined reporter gene locus allows studying the effects of MOF recruitment in male flies, where MOF is part of the dosage compensation system, and in females, where the function of MOF has not been studied yet.

We expressed MOF-Gal4 from the endogenous *mof* promoter to avoid distortions due to overexpression (see Figure S1A available online). The MOF-Gal4 transgene was functional, since it rescued the male-specific lethality of the *mof*¹ allele (Hilfiker et al., 1997). The fusion protein incorporated into the DCC, as Gal4 staining colocalized with the DCC on the male X chromosome as well as on a few autosomal sites (Figure S1D). Expression of MOF-Gal4 did not affect the survival or fertility of females. The RC5 element integrated at 93B (Zink and Paro, 1995) was stained with the Gal4 antibody in polytene chromosomes of female and male larvae (Figures S1C, S1D, and S2A–S2D). Binding of MOF-Gal4 to the UAS recruited the other DCC members in males, but not in females (Figures S2A–S2D).

Taken together, these results establish the functionality of MOF-Gal4 in our experimental setup, which enabled us to study its impact on reporter gene transcription in the different molecular contexts of male and female flies.

Different Modes of Activation by MOF in Male and Female Flies

We tested the ability of tethered MOF to activate transcription of the *lacZ* and *mini-white* genes in the context of the RC5 construct by quantitative RT-PCR analysis of RNA from sorted male and female adult flies. Expression of MOF-Gal4 activated both genes in both sexes. In males, the activation of the *lacZ* gene was in the 2-fold range relative to the level in the absence of MOF. Unexpectedly, the gene was expressed considerably stronger in females than in males (Figure 1C). We were concerned that assembly of a DCC at the UAS of RC5 in males might sterically hinder preinitiation complex formation at the minimal *hsp70* promoter of the *lacZ* gene, resulting in lower expression. Therefore, we generated the RC3 fly line in which the UAS is placed 3' of the *lacZ* gene (Figure 1B). Recruitment of MOF-Gal4 to the 3' UAS led to activation of *lacZ* in the 2-fold range in males, but stronger stimulation in females (Figure 1C). The *mini-white* gene, whose start site is about 4.5 kb away from the tethering site in either construct, was also induced by MOF-Gal4, with stronger effects in females (Figure 1D). Plotting the degree of activation of the reporter genes as a function of distance from the UAS to the promoters shows that recruitment of MOF leads to an approximately 2-fold activation of transcription within the tested zone of 5 kb in males. In contrast, activation in females was strongly distance dependent and stronger if MOF was close to the transcription start (Figure S1B). Taken together, these results document the activation potential of MOF in both sexes and suggest different modes of activation. The 2-fold, distance-independent activation in males is reminiscent of

a dosage compensation regime. However, in females the distance-sensitive activation that exceeds 2-fold is more akin to promoter-proximal activation.

As activation by MOF is thought to be based on its acetylating function, we wondered whether different levels or distribution patterns of the H4K16ac mark could explain the sex-specific activation differences. We adapted a chromatin immunoprecipitation (ChIP) protocol for adult flies (Negre et al., 2006), which allowed testing for the presence of regulators and the H4K16ac mark in hand-sorted male and female flies along the reporter using four PCR amplicons (Figures 2A–2E). Normalizing the ChIP values to background levels of MOF binding at a region 5' of the *GAPDH1* gene enabled a comparison of different chromatin preparations. The X chromosomal, dosage-compensated *armadillo* (*arm*) locus served as internal control for physiological, male-specific interaction of the DCC and H4K16 acetylation.

ChIP directed against MOF confirmed the recruitment of similar levels of MOF-Gal4 to the tethering site in both reporter loci and sexes (Figures 2A and 2B). Crosslinking of MOF was reduced with increasing distance from the tethering site, but still significant at the 3' end of the *mini-white* gene, 8.5 kb away from the UAS (compare to levels at the *arm* locus), consistent with the observed spreading of DCC from engineered autosomal HAS (Alekseyenko et al., 2008). MOF recruitment was similar in both reporter lines, although in RC5 the distance-dependent decline was more pronounced in females. The H4K16ac levels corresponded well to those of MOF-Gal4, with strong enrichments at the site of tethering and a distance-dependent reduction. In males, H4K16ac levels tended to spread further from the tethering site (Figures 2C and 2D). In the absence of MOF-Gal4 only background levels of endogenous MOF were observable.

MSL2 was recruited to the RC3 locus by MOF-Gal4 only in males, as expected since MSL2 is the male-specific determinant of the DCC (Figure 2E). Binding across the locus closely paralleled that of MOF. The recruitment of MSL2 confirmed that MOF functions in the context of a DCC in males, as suggested by the polytene chromosome staining. Strikingly, in the absence of the DCC in females, MOF catalyzed comparable levels of H4K16ac at the reporter locus. Unexpectedly, H4K16ac levels at the reporter genes do not correlate linearly with the corresponding transcription levels, as H4K16ac is moderately higher in males than in females (Figures 2C and 2D), yet the transcription is significantly stronger in females (Figures 1C and 1D). By contrast, transcription output correlated well with the number of RNA polymerase molecules at the 5' of the *lacZ* gene, determined by ChIP of the integral Rpb3 subunit (Muse et al., 2007) (Figure 2G). The levels of two other factors related to dosage compensated chromatin, the nucleosome-remodeling ATPase ISWI (Deuring et al., 2000) and the supercoiling factor SCF (Furuhashi et al., 2006) were unchanged upon recruitment of MOF (data not shown).

H4K16 acetylation appears not to be an absolute requirement for transcription since recruitment of an intact yeast GAL4 activator (without MOF) to the RC5 reporter led to profound activation to similar levels in both sexes without enrichment of H4K16ac (Figures S2E and S2F).

We conclude that the contributions of H4K16ac to transcription are highly context dependent. To further substantiate this

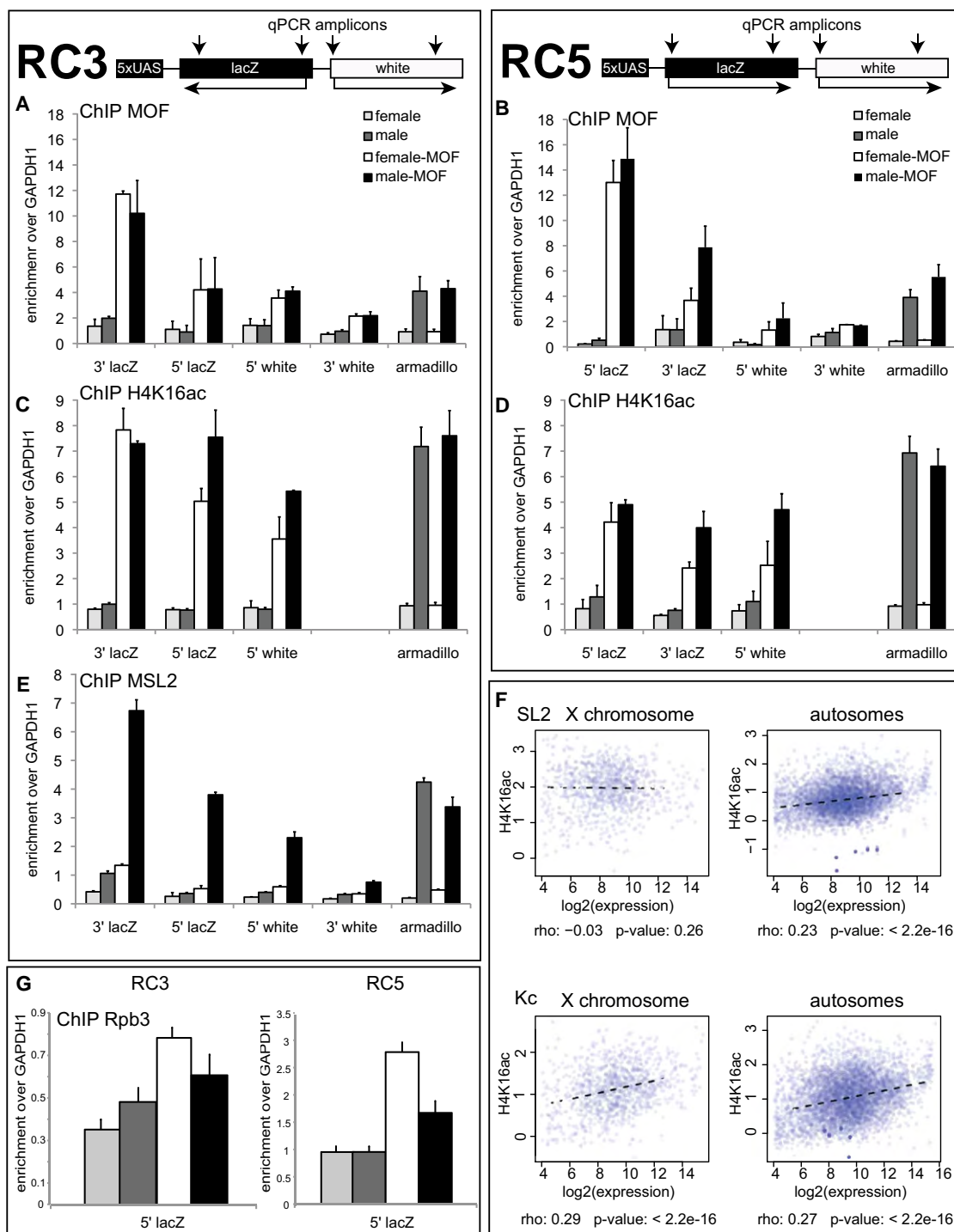


Figure 2. Chromatin Interactions upon MOF Recruitment

Chromatin constituents were monitored by ChIP in sorted male and female adult flies. The presence of sequences in the immunoprecipitate corresponding to four amplicons along the reporter loci (arrows, see schematics on top) was determined by qPCR. The RC3 locus was probed in (A), (C), and (E) and the RC5 locus in (B) and (D). Chromatin was from females or males of the reporter lines in the absence of MOF-Gal4 (light and dark gray bars, respectively) or in the presence of MOF-Gal4 (white and black bars). Error bars represent mean \pm SEM. (A) and (B) display MOF interactions at the RC3 and RC5 reporters, respectively. (C) and (D) show the corresponding H4K16 acetylation levels. (E) MSL2 interactions were revealed at the RC3 reporter. (F) Genome-wide correlation of gene expression and H4K16ac levels in SL2 cells, using the values of Kind et al. (2008). (G) Binding of the polymerase II subunit Rpb3 5' to the lacZ gene. All ChIP experiments were internally controlled by monitoring interactions at the X chromosomal and dosage compensated *armadillo* gene. Error bars represent mean \pm SEM. See also Figure S2.

conclusion, we performed an analysis of the global relationship between H4K16ac levels (Kind et al., 2008) and gene transcription in *Drosophila* SL2 cells. We found that H4K16ac correlates positively with transcription on all female chromosomes as well as on the male autosomes, but this correlation was not observed on the male X chromosome (Figure 2F). Taken together with our earlier finding that MOF can strongly activate transcription in yeast (Akhtar and Becker, 2000), the data suggest that H4K16 acetylation has a potential for strong transcriptional activation, which is diminished in the context of the DCC in males where a higher than 2-fold activation would lead to nonphysiological overcompensation.

Identification of MOF Interactors in Female Cells

So far the data suggest that MOF can activate transcription in two distinct settings. In male flies, MOF mainly resides in the DCC. The molecular context in females is not known. Akhtar and colleagues recently reported the purification of an alternative MOF complex (the NSL complex) from mixed-sex *Drosophila* embryos and the Schneider-derived Sf4 cell line, which has male features (Mendjan et al., 2006). In order to explore the existence of an alternative complex in female cells, we established a Kc cell population stably expressing FLAG-tagged MOF-Gal4. Control cells expressed only the hygromycin resistance, but no tagged protein. Extracts from both cell populations were subjected to FLAG-affinity purification, and MOF interactors were identified by mass spectrometry. Comparison of the identified peptides revealed robust scores for MOF, MBD-R2, NSL1, NSL2, NSL3, WDS, and MCERS2 exclusively in the MOF purification (Table S1A). These proteins had previously been identified as MOF interactors by Mendjan et al. (2006). The proteins Z4, Chriz/chromator, and exosome subunits that had also been suggested to interact with MOF were retrieved in our control purification and hence do not qualify as MOF associated (Table S1). In further contrast to the complex described by Mendjan et al., we did not find nuclear pore complex components in either purification. To rule out the possibility that substoichiometric amounts of nucleoporins escaped mass spectrometry detection, we probed for MTOR and Nup153 by western blotting but failed to detect them in our purifications (data not shown). Our data highlight MBD-R2, NSL1, NSL2, NSL3, WDS, and MCERS2 as subunits of one or more MOF complexes in female cells, which may provide the molecular context for MOF function.

MBD-R2 Is a Marker of Active Genes

The alternative MOF complex was purified from female cells, but the data of Mendjan et al. (2006) suggest that a related complex also exists in male cells. Since MBD-R2 was the MOF interactor with the highest score in both purifications, we provisionally term the alternative complex the “MOF-MBD-R2” complex. As we wished to follow MBD-R2 as a marker for this complex, we raised polyclonal antibodies against MBD-R2 and confirmed their specificity (Figures S3F–S3H). Staining of polytene chromosomes revealed localization of MBD-R2 at the majority of gene-rich interbands of all larval polytene chromosomes of both sexes (data not shown).

For high-resolution mapping, we combined ChIP with probing tiled microarrays (ChIP-chip) (Straub et al., 2008). To distinguish

between the sexes, we established chromosome-wide binding profiles in sex-sorted adult flies (Figures S3B–S3E and S6). We found genome-wide binding of MBD-R2 in female (Figure S3E) and male flies (Figure 3A) with a preference for genes over intergenic sequences (Figure S3A). In order to correlate the MBD-R2 chromosome interaction profile with gene expression, we generated the ChIP-chip profile of MBD-R2 in SL2 cells. In brief, the majority of binding events observed in adult flies also occur in SL2 cells (Figure S6B). We determined the transcriptional status of those cells by Affymetrix gene expression profiling of two biological replicates. We found that most MBD-R2-bound genes are transcriptionally active, whereas unbound genes are rarely transcribed (Figure 3B). MBD-R2, therefore, appears to be a marker of active chromatin. Interestingly, the amount of MBD-R2 loading is not directly proportional to the expression level. Above a certain (low) expression threshold, MBD-R2 binding does not systematically increase with transcription (Figure 3B). MBD-R2 binds the same loci in both sexes without significant preference for any of the chromosomes (Figures 3C and 3D). The male and female X chromosomes are bound to roughly similar extents (compare Figure 3A and Figure S3E, Figures 3C, 3D, 7B).

Colocalization of MOF with DCC and MBD-R2

To correlate MOF binding with either MBD-R2 or MSL1 (a marker for the DCC), we generated binding profiles for MOF in flies of sorted sex and for MSL1 in males. In agreement with previously published SL2 cell profiles (Gilfillan et al., 2006), MSL1 associates almost exclusively with X chromosomal genes in male flies (Figure S3D). Here, MOF predominantly colocalizes with MSL1. MOF binds the autosomes in males with much lower but still significant levels (Figure S3B), in agreement with mapping data in SL2 cells and polytene chromosome staining (Bhadra et al., 1999; Kind et al., 2008). By contrast, the MOF levels in female flies were similar on all chromosomes and clearly elevated relative to male autosomes (Figure S3C, Figure 7A).

MOF binding generally correlates with both MSL1 and MBD-R2 with interesting difference upon closer look. In the absence of the DCC, on male autosomes and all female chromosomes, MOF colocalizes extensively with MBD-R2 (correlation coefficient above 0.9, Figures 4A and 4B). Unexpectedly, most of the MOF-bound genes on the male X chromosome recruit both, MSL1 and MBD-R2 (Figure 4C). However, MOF correlates highly with MSL1, whereas the association of MBD-R2 varies (Figure 4A). An assessment of the binding profiles along the gene bodies shows that the distributions differ (Figure 4D). Next, we calculated the average binding signal along a window of 4 kb around the transcriptional start site (TSS) or the transcriptional termination (TT) site (Figure 4E). Avoiding complications arising from close-by, nested, and overlapping genes in the compact *Drosophila* genome, we selected nonoverlapping, active genes with a minimal length of 2 kb. To avoid bias by outliers, the genes were grouped according to their expression levels. The study revealed an enrichment of MSL1 (Figure S4) and MOF toward the 3' end of genes on the male X chromosome (Figure 4E), consistent with previous observations in SL2 cells (Gilfillan et al., 2006; Kind et al., 2008). In striking contrast to MOF and MSL1, MBD-R2 is enriched toward the 5' end of genes

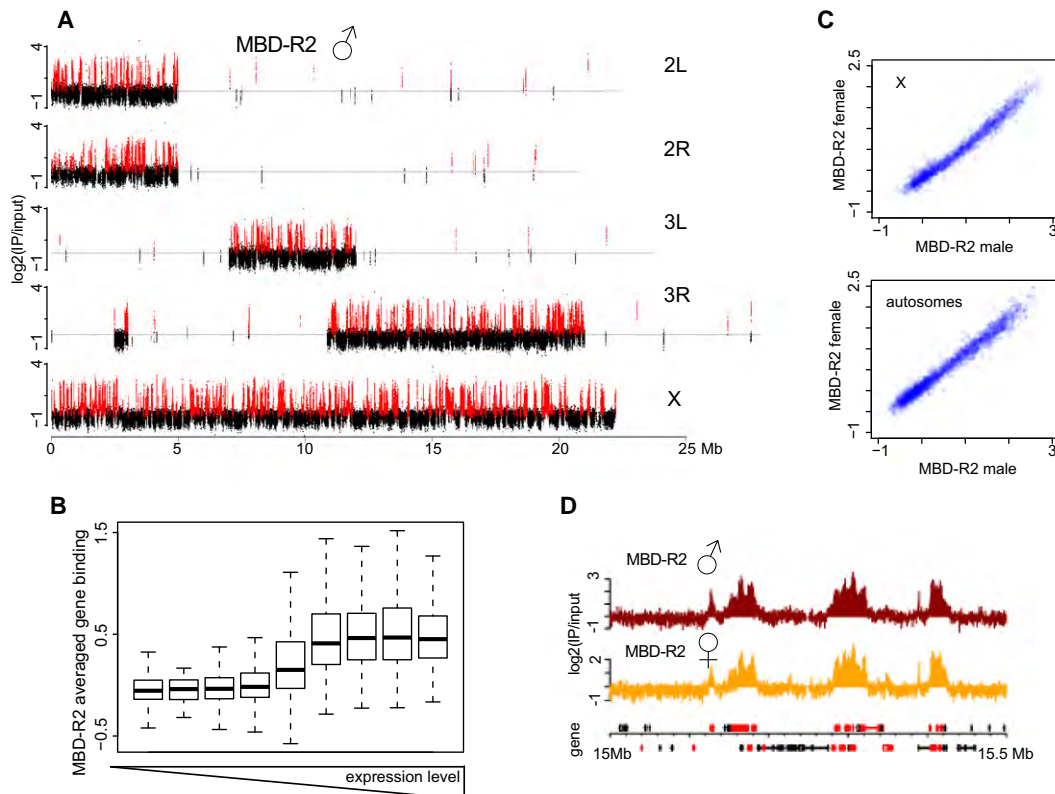


Figure 3. MBD-R2 Binds Active Genes Globally

(A) Chromosome-wide binding of MBD-R2 in male flies. Oligonucleotide probe signals are plotted along the major chromosomes as indicated (note that the entire X chromosome is represented, but only parts of each autosome). Significant bound probes (tileHMM, see the [Supplemental Experimental Procedures](#)) are marked in red. Numbers along the x axis denote the physical position along the chromosomes in megabase pairs (Mb).

(B) MBD-R2 binds active genes independent of expression level. Genes were grouped in equal-sized bins according to their expression levels (GST RNAi expression set, see [Figure 5](#)). The ordinate depicts the signal distribution of average gene binding values in the respective groups, assessed by MBD-R2 ChIP-chip in SL2 cells.

(C) MBD-R2 binding coincides in male and female flies. Correlation of the ChIP-chip binding signal (averaged gene binding score, see the [Supplemental Experimental Procedures](#)) of X chromosomal genes (upper panel) or autosomes (lower panel) from at least three biological replicates from both sexes. Spearman correlation coefficient $\rho = 0.94$ ($p < 2.2 \times 10^{-16}$).

(D) ChIP-chip profiles of MBD-R2 in male and female flies along a representative X chromosomal region. The profiles are related to a gene representation at the bottom of the panel, where genes drawn above the line are transcribed from left to right and genes below the line are transcribed from right to left. Active genes are marked in red, inactive genes in black (Chintapalli et al., 2007). The x and y axis are denoted as in (A).

See also [Figure S3](#).

and peaks around the TSS ([Figure 4E](#)). In the absence of the DCC, on all female genes and on male autosomal genes, MOF enriches with MBD-R2 at the 5' ends.

In summary, we generated a high-resolution chromosome-wide binding profile of MOF, MSL1, and MBD-R2 in primary fly tissue (summarized in [Figure S6B](#)). The DCC-binding pattern is very similar in male flies and in SL2 cells. Colocalization suggests that MOF resides in the DCC on the male X chromosome and with the MOF-MBD-R2 complex on all other chromosomes. Most active genes are bound by the MOF-MBD-R2 complex, which enriches at the 5' end of genes in the absence of DCC.

MBD-R2 Acts as a Genome-wide Transcriptional Activator

MBD-R2 preferentially associates with the 5' ends of active genes. Does it contribute to transcriptional activation? Unfortun-

nately, flies bearing mutant alleles of *MBD-R2* die early during development precluding a loss-of-function analysis. We thus ablated MBD-R2 in SL2 cells by RNA interference (RNAi) ([Figure S3G](#)) and investigated the transcriptome changes. Transcription profiles were obtained from two biological replicates of cells subjected to RNAi targeting the central and the 3' part of the MBD-R2 transcript and controlled for nonspecific effects with RNAi directed at glutathion-S-transferase (GST) sequences. The transcriptomes of the two MBD-R2 RNAi samples and of the biological replicates were highly similar, precluding off-target effects ([Figure S5](#)). Next, we sorted all active genes in equal-sized bins according to the MBD-R2 ChIP-chip signal. Plotting those bins against the expression change after RNAi revealed that the more a given gene was bound by MBD-R2, the more its expression was reduced after ablation of the factor ([Figure 5](#)). There was no significant difference between X and autosomes in

this respect (data not shown). We conclude that MBD-R2 is a transcription activator.

The DCC Integrates Activating and Restrictive Principles

The enrichment of MBD-R2 toward the 5' ends of genes and its global stimulation of transcription suggests a contribution of MBD-R2 to the activation of the reporter loci. In support of this hypothesis, we found by ChIP roughly twice as much MBD-R2 at the 5' end of the lacZ gene upon MOF recruitment in female flies compared to males (Figure 6D). Figure 6A graphically illustrates the relative occupancy of MBD-R2, MSL2, and tethered MOF (H4K16 acetylation) at the 5' lacZ site in RC5 males and females (data from Figures 2 and 6D and not shown). MOF recruitment and H4K16ac levels are similar in both sexes. In females lacking MSL2, MBD-R2 accumulates to higher levels than in males. Conceivably, here all MOF is associated with MBD-R2. The situation is more complicated in male cells, where MSL2 (the DCC) localizes to the tethering site in addition to MBD-R2. Given the limited number of UAS elements, it is possible that MBD-R2 and the MSL proteins compete for interaction with tethered MOF. In order to test this hypothesis, we expressed MSL2 in females bearing the RC5 reporter, which leads to ectopic formation of the DCC and transforms the flies into pseudomales (Kelley et al., 1995). MSL2 was recruited to the tethering site by ChIP, suggesting the assembly of the DCC (Figure 6B). Strikingly, at the same time the levels of MBD-R2 at this site were reduced to about 50%, in support of a competition between MSL2 and MBD-R2 for MOF binding (Figure 6D). Remarkably, measuring β -galactosidase activity under those conditions revealed a reduction of reporter gene expression, down to roughly 2-fold above basal levels (Figure 6E). Of note, MOF and H4K16ac levels did not change upon MSL2 expression (data not shown). This result documents the lower transcription activation potential of the DCC relative to the MBD-R2 complex at the reporter locus. We tested whether elevation of another DCC subunit, MSL1, also constrained MOF-dependent activation. Overexpression of MSL1 in the context of MOF-Gal4 recruitment in females does not lead to formation of the DCC (males show reduced viability) (Chang and Kuroda, 1998). Recruitment of MSL1 by MOF (Figure 6C) only led to a slight reduction of MBD-R2 levels (Figure 6D), concomitant with a minor reduction of lacZ expression (Figure 6F).

The reduced expression of the reporter locus upon DCC assembly in females led us to hypothesize that the DCC may not only contain an activator (MOF) but also a "constraining principle" that limits the overall activation to a 2-fold range. In order to test this hypothesis, we resorted to SL2 cells, which contain both the DCC and MBD-R2 complex. We cotransfected these cells with an expression plasmid for MOF-Gal4 and plasmids bearing UAS_{Gal}-responsive luciferase reporter genes whose transcription we monitored. The UAS_{Gal} tethering site was placed either 5' or 3' of the reporter gene in an effort to create a situation resembling that of RC5 and RC3, respectively. In order to selectively monitor the function of either complex alone, we knocked down either MSL2 or MBD-R2 using two nonoverlapping RNAi constructs. Remarkably, the luciferase expression from both constructs increased relative to the control RNAi upon

ablation of MSL2, in support of the hypothesis that the DCC restricts the activation potential of MOF activation (Figure 6G). The increased activation may well be due to increased association of MBD-R2 with tethered MOF, since RNAi against MBD-R2 led to a reduction of reporter gene expression (Figure 6G).

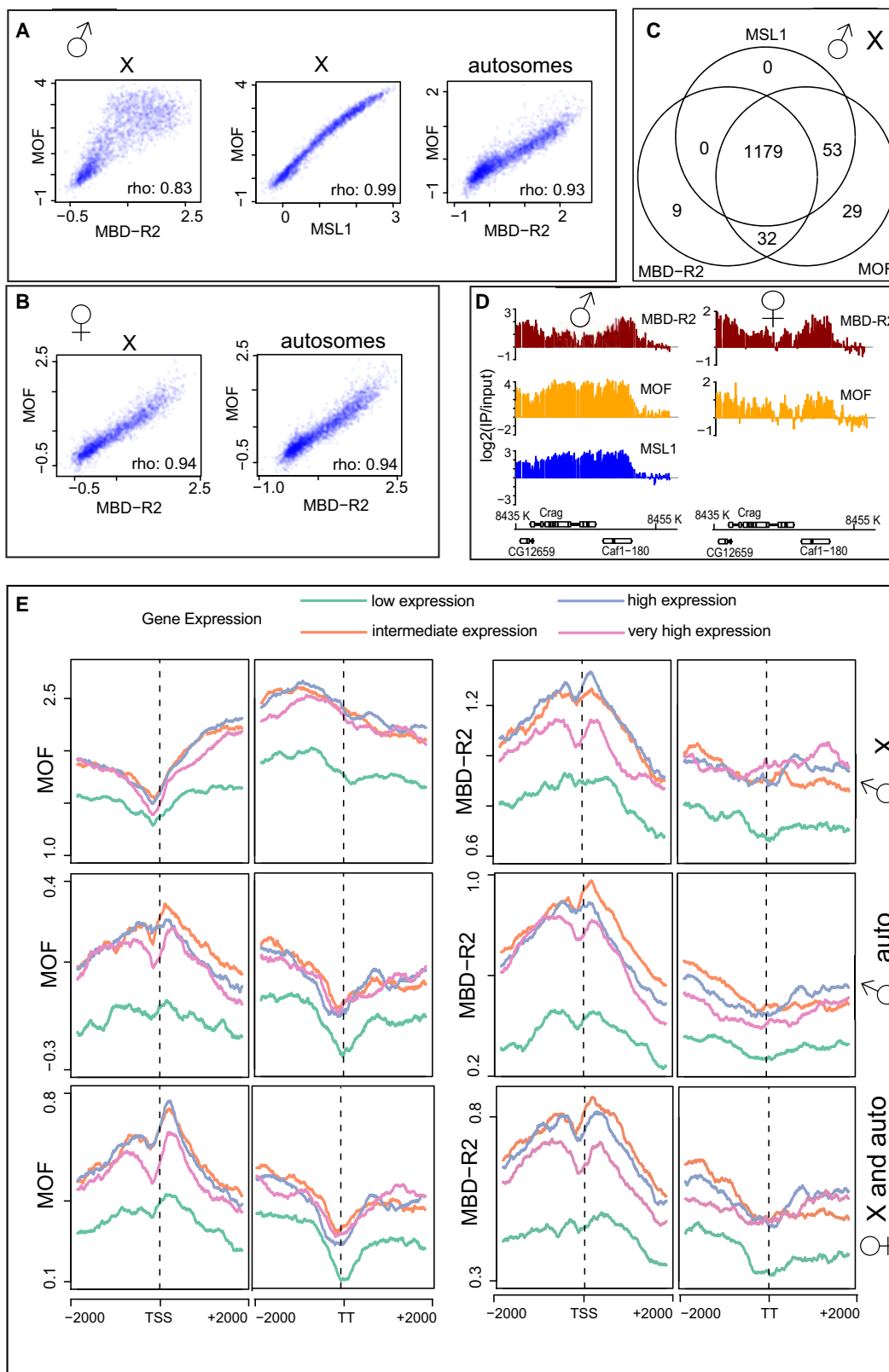
Dynamic Distribution of MOF between the DCC and the MBD-R2 Complex

These data suggest a dynamic distribution of MOF between DCC and MBD-R2 complexes, which depends on the relative levels of either component. In order to test whether such a scenario would also apply on a genome-wide scale, we generated "pseudomale" flies as before by ectopic expression of MSL2 in females and generated chromosome-wide binding profiles for MBD-R2 and MOF. Given the differential distribution of MOF in male and female cells (Figures S3B and S3C), we expected to see a relocalization of MOF from the autosomes to the X chromosome. Strikingly, we found this to be the case. Even modest expression of MSL2 (and concomitant DCC assembly) led to a significant relocalization of MOF from autosomes to the X chromosome (Figures 7A and 7C), while the localization of MBD-R2 did not change (Figure 7B). This result suggests that the relative levels of the two targeting principles, MBD-R2 and DCC, determine the distribution of MOF between the X and the autosomes.

DISCUSSION

Acetylation of H4K16ac is unique among the histone acetylation events in that it reduces the interaction between H4 and H2A/H2B in close-by nucleosomes and hence counteracts the compaction of nucleosomal arrays into fibers (Robinson et al., 2008; Shogren-Knaak et al., 2006). The large potential of H4K16ac and of MOF, the enzyme responsible for its placement, for regulated chromatin opening and activation of transcription has been revealed in model systems (Akhtar and Becker, 2000). MOF was originally discovered in *Drosophila* due to its vital role in dosage compensation, which so far is mainly attributed to its H4K16 acetylation function (Akhtar and Becker, 2000; Hilfiker et al., 1997; Smith et al., 2000). Dosage compensation involves a subtle, 2-fold activation of transcription of most X chromosomal genes in males. The principles that allow such precise adjustment of transcription are unknown and the contribution of H4K16 acetylation to this fine-tuning is equally mysterious.

Despite the male-specific lethality associated with MOF mutations, MOF also has less prominent functions in female cells, where it is expressed at considerable levels (Bhadra et al., 2000; Kind et al., 2008). We have now identified an additional molecular context for MOF function in female Kc cells by identifying MOF interactors. Previously, Akhtar and colleagues had shown the existence of a highly related complex in male cells and mixed sex embryos (Mendjan et al., 2006). A hallmark of this complex is the subunit MBD-R2, an uncharacterized protein featuring similarity to methyl-CpG-binding domains and several types of zinc fingers (Bienz, 2006). Because of the differences between the interactors we purified and the NSL complex described by Mendjan et al. (2006), we provisionally refer to the MOF complex in Kc cells as the MOF-MBD-R2 complex.



The stoichiometry of its components or the homogeneity of assemblies they form is not known yet.

Male cells contain MBD-R2 (or NSL) complexes in addition to the DCC, and Mendjan et al. (2006) suggested that the NSL complex and the DCC might compete for MOF recruitment. Biochemical studies are needed to firmly establish such a scenario. However, our chromosome-wide profiles are indeed consistent with a dynamic distribution of MOF between these two complexes. The global chromosomal interaction pattern of MOF correlates well with the association profile of MBD-R2; conceivably, the MBD-R2 context is the only targeting determinant for MOF in female cells. This targeting scenario is challenged by forced expression of MSL2 in females that leads to ectopic DCC assembly in potential competition for MOF. Remarkably, we observed a massive relocalization of MOF from the autosomes to the X chromosome. This creates a situation reminiscent of male cells, where we find that MOF preferentially colocalizes with MBD-R2 on autosomes, but with the DCC on the X chromosome. We suppose that the balancing of the genome—dosage compensation—involves a dynamic distribution of MOF between the two complexes.

Our data suggest that the two MOF complexes have different roles in transcription. The DCC mainly localizes to gene bodies with a 3' enrichment, which has led to the speculation that dosage compensation regulates elongation efficiency (Aleksyenko et al., 2006; Gilfillan et al., 2006). MBD-R2 also activates transcription but, in contrast to the DCC, binds active genes globally with a 5' bias. The defined setting of our reporter gene system revealed that, depending on context, tethering MOF installed two different activation modes. In females, in the context of MBD-R2, activation was variable, distance dependent, and correlated with H4K16 acetylation levels as well as enhanced polymerase loading. This scenario is akin to short-range stimulation of transcription initiation by “promoter-specific” regulators.

By contrast, transcription stimulation in males was dominated by the DCC. Activation was in the 2-fold range and distance independent, a scenario compatible with dosage compensation, despite of the high H4K16 acetylation levels that correlate with higher transcription rates in female cells. We thus hypothesized that the activating effect of H4K16 acetylation is constrained by an opposing principle associated with the formation of the DCC. This hypothesis has been strengthened by several of our results. First, a global analysis reveals that H4K16ac levels

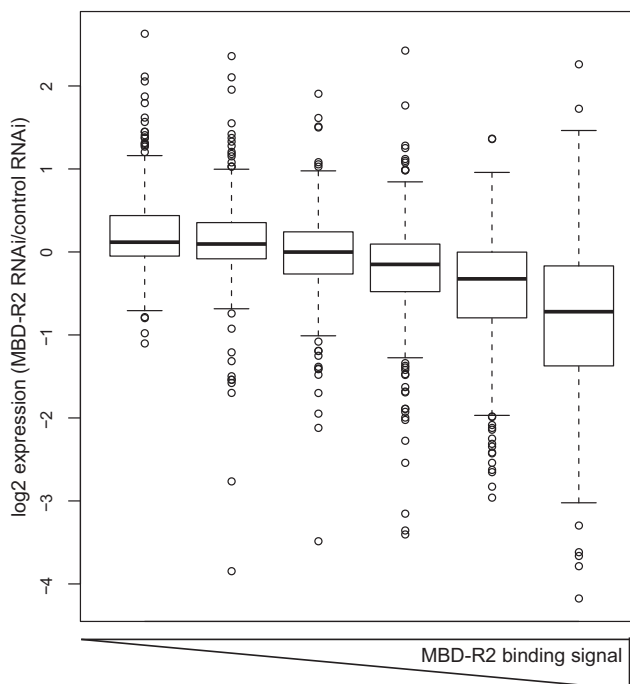


Figure 5. MBD-R2 Is a Genome-wide Transcriptional Activator

Active genes (Muse et al., 2007) were grouped according to their averaged MBD-R2 binding score and plotted against their expression change after MBD-R2 RNAi (\log_2 [MBD-R2 RNAi/control RNAi]). Genes were categorized in groups with increasing binding score as represented in the box plots. Genes were considered “active” if elongating polymerase was bound significantly in the ChIP-chip profile of Muse et al. (2007). Similar results were obtained if genes with an expression value of at least 4 in the control samples (Affymetrix scale) were classified as “active.”

See also Figure S5.

roughly correlate with transcription on all female chromosomes and on male autosomes, but not on the male X chromosome. Second, reduction of MSL2 levels in SL2 cells led to activation of a reporter gene, while ablation of MBD-R2 resulted in reduced transcription. Finally, ectopic expression of MSL2 in females resulted in reduced levels of MBD-R2 at the reporter locus and limited the stimulation of transcription to a 2-fold range.

The finding that a HAT may be programmed by association with distinct sets of subunits is not without precedent. For example, the catalytic activity of GCN5 is modulated depending

Figure 4. MBD-R2 and MOF Colocalize on Active Genes under All Circumstances, Except for the Male X Chromosome

(A and B) Correlation between chromosomal binding of MOF and MSL1 or MOF and MBD-R2 on male (A) or female (B) chromosomes. Scatter plots of averaged binding score per gene (see the Supplemental Experimental Procedures) of MBD-R2, MOF, and MSL1 on male chromosomes as indicated. All Spearman correlation coefficients (ρ) are $p < 2.2 \times 10^{-16}$.

(C) The majority of active genes on the male X are bound by MBD-R2, MOF, and MSL1. Venn diagram of significantly bound active genes.

(D) Binding of MSL1, MOF, and MBD-R2 along a representative region of the X chromosome in males (left panel) and females (right panel). The profiles are related to the gene representation at the bottom of the panel. All genes are active. Genes depicted above or below the line of physical coordinates are transcribed from left to right or right to left, respectively.

(E) Differential distribution of MOF and MBD-R2 across the gene bodies. Cumulative plots of MOF (left) and MBD-R2 (right). ChIP-binding signals along genes were aligned at the transcriptional start site (TSS) and transcriptional termination site (TT). The average binding score is plotted 2 kb bp up- and downstream of the TSS and TT in a sliding window (10 bp step size, 300 bp window). For this analysis, 637 X chromosomal genes and 740 autosomal genes were selected, which are minimally 2 kb long, do not overlap with other genes, and are expressed. The genes were grouped according to their gene expression level (Chintapalli et al., 2007) into “low” (black), “intermediate” (orange), “high” (blue), and “very high” (purple) categories. The x axis denotes the distance to the TSS and TT in base pairs. See also Figure S4.

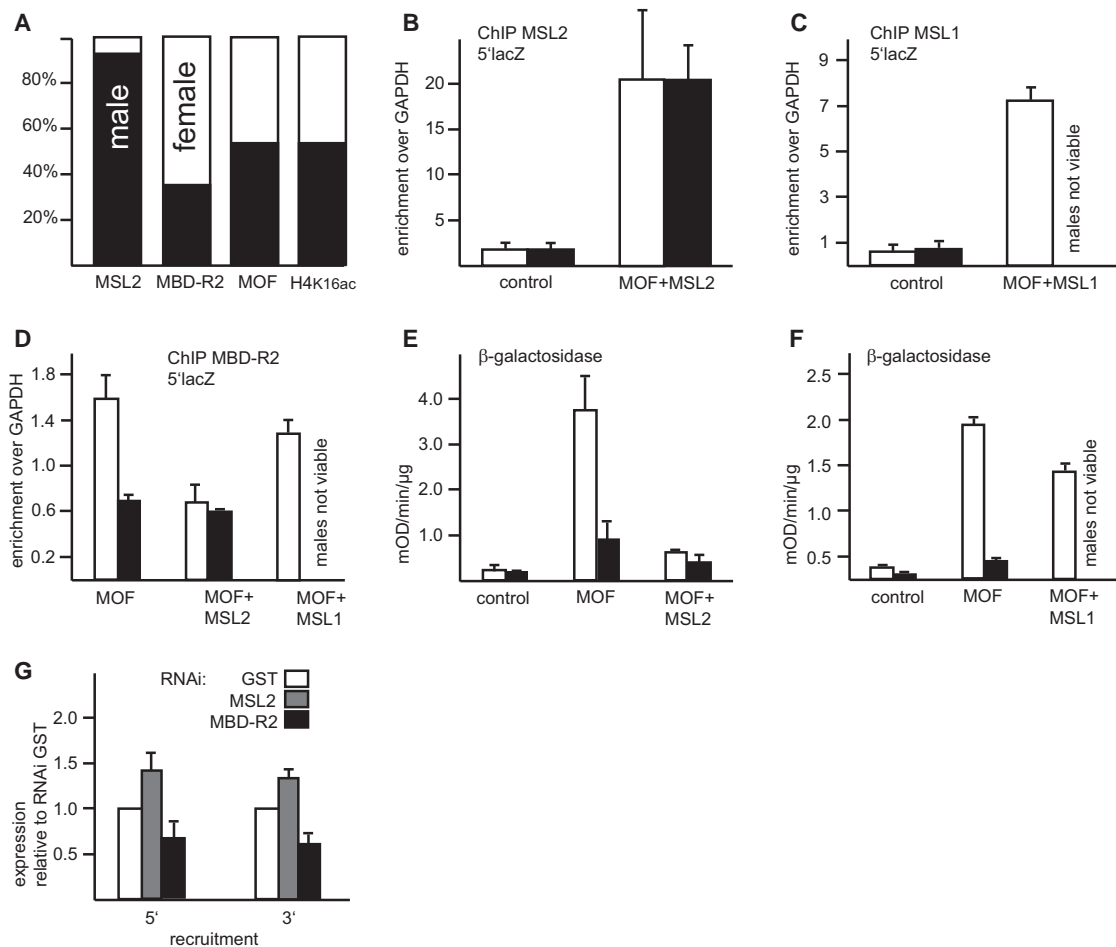


Figure 6. Sex-Specific Cofactor Recruitment upon MOF Binding

(A) Schematic illustration of the relative occupancy of chromatin constituents at the 5' lacZ site in male and females. The graph summarizes data presented in Figures 2B, 2D, and 6B and data not shown for MSL2. Displayed are the ratios of ChIP enrichment of the indicated antibodies in females (white bar) and males (black bar). MOF or H4K16ac levels are similar in both sexes. MSL2 is only enriched in males (background levels in females). Females lacking MSL2 accumulate more MBD-R2 than males.

(B and C) Ectopic expression of MSL2 (B) and MSL1 (C) in females leads to enrichment of the corresponding protein at the RC5 locus upon MOF recruitment (control represents the RC5 locus in the absence of MOF-Gal4). Error bars represent mean \pm SEM. Overexpressing MSL1 males could not be analyzed due to lethality. White bars, female flies; black bar, male flies.

(D) ChIP with MBD-R2 in females expressing either MSL2 as in (B) or MSL1 (C). Tethering MOF to the lacZ reporter leads to a 2-fold enrichment of MBD-R2 in females versus males. Expression of MSL2 (pseudomales) reduces the MBD-R2 levels. Expression of MSL1 has only a minor impact. Presentation as in (B).

(E and F) Change of lacZ induction (β -galactosidase activity) by MOF upon ectopic expression of MSL2 (E) or MSL1 (F). Presentation as in (B).

(G) Reporter gene activity in SL2 cells after ablation of MSL2 and MBD-R2. SL2 cells were treated with dsRNA against GST (control), MSL2, or MBD-R2 sequences. Three days after RNAi treatment, cells were transfected with an expression vector for Gal4-MOF and a firefly luciferase reporter furnished with UAS_{Gal} elements either 5' of the minimal thymidine kinase promoter or 3' of the firefly luciferase gene. Transfection efficiency was normalized by coexpression of a Renilla luciferase expression plasmid. T test, two-sided and unpaired, error bars represent mean \pm SEM.

on the SAGA or ATAC context (Carre et al., 2008). While this manuscript was revised, Conaway and colleagues reported on a change in substrate specificity for human MOF, depending on its residence within the MSL or NSL context (Cai et al., 2010). It will be interesting to see whether the different complex environments also reprogram the catalytic activity of MOF in *Drosophila*.

In summary, our data suggest that the precise 2-fold activation in *Drosophila* dosage compensation is achieved by constraining the activation potential of H4K16 acetylation. Deciphering the underlying mechanism will help to reveal the principles that

govern genome balancing at the level of chromosome organization and remains as a challenge for future work.

EXPERIMENTAL PROCEDURES

Further detailed experimental procedures are described in the [Supplemental Information](#).

Fly Stocks

All *P* element constructs were injected into *y¹,w¹¹¹⁸* and raised on standard fly food.

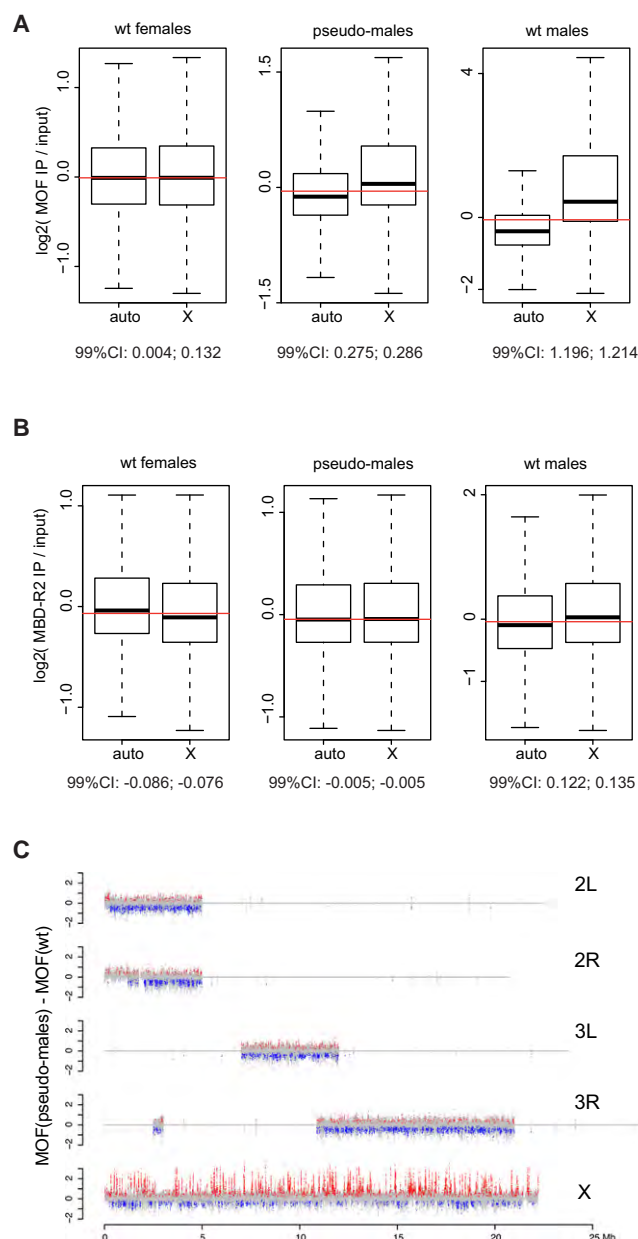


Figure 7. MOF Redistributes to the X Chromosome in Female Flies Overexpressing MSL2

(A and B) Box plots present ChIP-chip signals of autosomal or X chromosomal probes, which represent the binding of MOF (A) and MBD-R2 (B) in wild-type (WT) females, MSL2 expressing females (pseudomales) and wild-type males. The ordinates depict the log₂-normalized values of probes. The 99% confidence interval (CI) of the true difference in means between X chromosomal and autosomal MOF binding is shown. The red line depicts the median from all probes.

(C) Chromosome-wide difference plots of MOF binding between pseudomales and females. The difference of oligonucleotide probe signals between pseudomales and females is plotted along the major chromosomes as indicated (note that only part of the autosomes are represented). Significant (adjusted p values < 0.05) gains are marked in red, significant losses are marked in blue. Numbers along the x axis denote the physical position along the chromosomes in megabase pairs (Mb).

See also Figure S6.

The RC5 reporter fly line is a gift from R. Paro and was originally published as U/I5 in Zink and Paro (1995). Fly strains expressing Hsp70Gal4, Hsp83MSL1, and Hsp70MSL2 are published (Chang and Kuroda, 1998; Rank et al., 2002; Straub et al., 2005b).

Antibodies

The MOF antibody was raised in rabbit using the complete protein as antigen (Akhtar and Becker, 2000; Morales et al., 2004). An MBD-R2 fragment corresponding to amino acids 1–355 was expressed in *E. coli* BL21 cells fused to an N-terminal GST tag, purified on glutathione beads, and used to raise polyclonal antibodies in rabbits. Antibodies against H4K16ac were obtained from ACTIVE MOTIF (#39167) and the GAL4 antibody from SantaCruz Biotechnology (sc-577). MSL1 and MSL2 antibodies were previously published (Gilfillan et al., 2006, 2007).

RNA Quantification

Total RNA was isolated from six adult females or seven males. Flies were frozen in liquid nitrogen, grained to powder, and suspended and purified with Trizol (Invitrogen). cDNA was synthesized with SuperScript II Reverse Transcriptase (Invitrogen). Quantification was performed with the real-time PCR ABI 7000 (primer see qPCR primers). Values were normalized to *gapdh1*.

ChIP-qPCR

ChIP material was quantified by real-time PCR (ABI 7000, Applied Biosystems). Approximately 1 ng of the purified input DNA was used per reaction. Depending on the antibody, 2%–5% of the ChIP material was required per reaction. Enrichment was calculated over input. Since various chromatin preparations of different genotypes and sexes were compared, we standardized to the autosomal *gapdh1* locus after input normalization. ChIP experiments were performed in technical and biological replicates. Primers used are indicated in Table S1. Each experiment was performed at least in triplicate.

ACCESSION NUMBERS

The microarray data are located at the Gene Expression Omnibus under the accession numbers GSE20695 (ChIP data) and GSE20744 (Affymetrix data).

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, four tables, Supplemental Experimental Procedures, and Supplemental References and can be found with this article at doi:10.1016/j.molcel.2010.05.022.

ACKNOWLEDGMENTS

We are grateful to the following colleagues: Axel Imhof for mass spectrometry, Angelika Mitterweger for help with ChIPs in SL2 cells, members of the Becker lab for helpful discussion, Mitzi Kuroda and Renato Paro for providing fly stocks, Karen Adelman for providing the Rpb3 antibody. This work was supported by the Deutsche Forschungsgemeinschaft through SFB-TR5 and the Gottfried-Wilhelm-Leibniz Programme. C.F. is a fellow of the International Max-Planck Research School in Munich. M.P. and P.B.B. designed the study. M.P. conceived and performed experiments. M.P. supervised C.F. C.F. conceived and performed genomic experiments. T.S. supervised C.F. in bioinformatic analyses. H.M. supported M.P. with fly genetics. M.P., C.F., and P.B.B. wrote the manuscript.

Received: August 26, 2009

Revised: January 12, 2010

Accepted: April 1, 2010

Published: June 24, 2010

REFERENCES

- Akhtar, A., and Becker, P.B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol. Cell* 5, 367–375.
- Alekseyenko, A.A., Larschan, E., Lai, W.R., Park, P.J., and Kuroda, M.I. (2006). High-resolution ChIP-chip analysis reveals that the *Drosophila* MSL complex selectively identifies active genes on the male X chromosome. *Genes Dev.* 20, 848–857.
- Alekseyenko, A.A., Peng, S., Larschan, E., Gorchakov, A.A., Lee, O.K., Kharchenko, P., McGrath, S.D., Wang, C.I., Mardis, E.R., Park, P.J., and Kuroda, M.I. (2008). A sequence motif within chromatin entry sites directs MSL establishment on the *Drosophila* X chromosome. *Cell* 134, 599–609.
- Bhadra, U., Pal-Bhadra, M., and Birchler, J.A. (1999). Role of the male specific lethal (msl) genes in modifying the effects of sex chromosomal dosage in *Drosophila*. *Genetics* 152, 249–268.
- Bhadra, U., Pal-Bhadra, M., and Birchler, J.A. (2000). Histone acetylation and gene expression analysis of sex lethal mutants in *Drosophila*. *Genetics* 155, 753–763.
- Bienz, M. (2006). The PHD finger, a nuclear protein-interaction domain. *Trends Biochem. Sci.* 31, 35–40.
- Cai, Y., Jin, J., Swanson, S.K., Cole, M.D., Choi, S.H., Florens, L., Washburn, M.P., Conaway, J.W., and Conaway, R.C. (2010). Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. *J. Biol. Chem.* 285, 4268–4272.
- Carre, C., Ciurciu, A., Komonyi, O., Jacquier, C., Fagegaltier, D., Pidoux, J., Tricoire, H., Tora, L., Boros, I.M., and Antoniewski, C. (2008). The *Drosophila* NURF remodelling and the ATAC histone acetylase complexes functionally interact and are required for global chromosome organization. *EMBO Rep.* 9, 187–192.
- Chang, K.A., and Kuroda, M.I. (1998). Modulation of MSL1 abundance in female *Drosophila* contributes to the sex specificity of dosage compensation. *Genetics* 150, 699–709.
- Chintapalli, V.R., Wang, J., and Dow, J.A. (2007). Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat. Genet.* 39, 715–720.
- Deuring, R., Fanti, L., Armstrong, J., Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S., Berloco, M., et al. (2000). The ISWI chromatin remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. *Mol. Cell* 5, 355–365.
- Fischle, W. (2008). Talk is cheap—cross-talk in establishment, maintenance, and readout of chromatin modifications. *Genes Dev.* 22, 3375–3382.
- Furuhashi, H., Nakajima, M., and Hirose, S. (2006). DNA supercoiling factor contributes to dosage compensation in *Drosophila*. *Development* 133, 4475–4483.
- Gelbart, M.E., and Kuroda, M.I. (2009). *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* 136, 1399–1410.
- Gelbart, M.E., Larschan, E., Peng, S., Park, P.J., and Kuroda, M.I. (2009). *Drosophila* MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. *Nat. Struct. Mol. Biol.* 16, 825–832.
- Gilfillan, G.D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P.B. (2006). Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex. *Genes Dev.* 20, 858–870.
- Gilfillan, G.D., Konig, C., Dahlsveen, I.K., Prakoura, N., Straub, T., Lamm, R., Fauth, T., and Becker, P.B. (2007). Cumulative contributions of weak DNA determinants to targeting the *Drosophila* dosage compensation complex. *Nucleic Acids Res.* 35, 3561–3572.
- Hamada, F.N., Park, P.J., Gordadze, P.R., and Kuroda, M.I. (2005). Global regulation of X chromosomal genes by the MSL complex in *Drosophila melanogaster*. *Genes Dev.* 19, 2289–2294.
- Hilfiker, A., Hilfiker, K.D., Pannuti, A., and Lucchesi, J.C. (1997). mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO J.* 16, 2054–2060.
- Kelley, R.L., Solovyeva, I., Lyman, L.M., Richman, R., Solovyev, V., and Kuroda, M.I. (1995). Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in *Drosophila*. *Cell* 81, 867–877.
- Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P., and Akhtar, A. (2008). Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in *Drosophila*. *Cell* 133, 813–828.
- Lucchesi, J.C., Kelly, W.G., and Panning, B. (2005). Chromatin remodeling in dosage compensation. *Annu. Rev. Genet.* 39, 615–651.
- Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., et al. (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol. Cell* 21, 811–823.
- Morales, V., Straub, T., Neumann, M.F., Mengus, G., Akhtar, A., and Becker, P.B. (2004). Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. *EMBO J.* 23, 2258–2268.
- Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. *Nat. Genet.* 39, 1507–1511.
- Negre, N., Hennein, J., Sun, L.V., Lavrov, S., Bellis, M., White, K.P., and Cavalli, G. (2006). Chromosomal distribution of PcG proteins during *Drosophila* development. *PLoS Biol.* 4, e170. 10.1371/journal.pbio.0040170.
- Qian, S., and Pirrotta, V. (1995). Dosage compensation of the *Drosophila* white gene requires both the X chromosome environment and multiple intragenic elements. *Genetics* 139, 733–744.
- Rank, G., Prestel, M., and Paro, R. (2002). Transcription through intergenic chromosomal memory elements of the *Drosophila* bithorax complex correlates with an epigenetic switch. *Mol. Cell. Biol.* 22, 8026–8034.
- Robinson, P.J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R.G., and Rhodes, D. (2008). 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *J. Mol. Biol.* 381, 816–825.
- Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844–847.
- Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., and Lucchesi, J.C. (2000). The *drosophila* MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. *Mol. Cell. Biol.* 20, 312–318.
- Straub, T., and Becker, P.B. (2007). Dosage compensation: the beginning and end of generalisation. *Nat. Rev. Genet.* 8, 47–57.
- Straub, T., Gilfillan, G.D., Maier, V.K., and Becker, P.B. (2005a). The *Drosophila* MSL complex activates the transcription of target genes. *Genes Dev.* 19, 2284–2288.
- Straub, T., Neumann, M.F., Prestel, M., Kremmer, E., Kaether, C., Haass, C., and Becker, P.B. (2005b). Stable chromosomal association of MSL2 defines a dosage-compensated nuclear compartment. *Chromosoma* 114, 352–364.
- Straub, T., Grimaud, C., Gilfillan, G.D., Mitterweger, A., and Becker, P.B. (2008). The chromosomal high-affinity binding sites for the *Drosophila* dosage compensation complex. *PLoS Genet.* 4, e1000302. 10.1371/journal.pgen.1000302.
- Zink, D., and Paro, R. (1995). *Drosophila* Polycomb-group regulated chromatin inhibits the accessibility of a trans-activator to its target DNA. *EMBO J.* 14, 5660–5671.

3.1.3 Supplementary data and figures

Molecular Cell, Volume 38

Supplemental Information

The Activation Potential of MOF Is Constrained for Dosage Compensation

**Matthias Prestel, Christian Feller, Tobias Straub,
Heike Mitlöhner, and Peter B. Becker**

Supplemental figures	page
Figure S1: supplement to Figure 1	(A) Western blot for determining transgenic MOF expression in adult flies (B) Sex specific reporter gene activation by MOF-Gal4 (C,D) Staining of polytene salivary gland chromosomes to visualize MOF-Gal4 incorporation and functionality 2
Figure S2: supplement to Figure 2	(A-D) Polytene stainings upon MOF recruitment in the RC5 line, zoomed in at the 93B integration site (E,F) The yeast GAL4 activator regulates gene expression independent on H4K16 acetylation 3
Figure S3: supplement to Figure 3	(A-E) Genome-wide binding of MBD-R2 and MOF in females and males, and MSL1 in males (F-H) MBD-R2 antibody specificity control 4
Figure S4: supplement to Figure 4	Cumulative profiles of MSL1 binding along genes 5
Figure S5: supplement to Figure 5	(A) Reproducibility of biological replicates of MBD-R2 ChIP-chip in males along a representative X chromosomal region (B) Overview of ChIP-chip profiles generated in this study: binding along a representative region on the X 6
Figure S6: supplement to Figure 7	Quality Control of biological replicates: Affymetrix transcriptome after MBD-R2 RNAi 7
Supplemental tables	
Table S1	MOF interactors in Kc cells 8
Table S2	Cloning Primer 9
Table S3	ChIP-qPCR primer 9
Table S4	RNAi primer 10
Supplemental experimental procedures	
Cloning procedures	11
Stable cell lines and protein purification	11
β -Galactosidase assay	12
Luciferase assay	13
Chromatin preparation and chromatin immunoprecipitation	13
ChIP-chip data analysis	14
Transcriptome analysis	15
Supplemental references	16

Supplemental figures

Figure S1: (A) Transgenic MOF expression in adult flies. The expression of the MOF fusion proteins in adult heads were tested in a RC5 background: females (lane 1 and 3) show less endogenous and also less transgenic MOF than males (lane 2 and 4). Lamin was used as a loading control.

(B) Sex-specific reporter gene activation by MOF-Gal4. Activation of the lacZ and mini-white gene upon recruitment of MOF-Gal4 was correlated to the distance between the UAS site and the promoters of the reporter genes. In females the fold activation decreases with increasing distance, whereas males maintain an approximately 2-fold activation.

(C,D) Staining of polytene salivary gland chromosomes to visualise MOF-Gal4 incorporation and functionality. Salivary glands of male and female larvae of the RC5 line were dissected, fixed, squashed and stained according to standard procedures (Dahlsveen et al. 2006). (B) Females show no binding of MSL3 but MOF-Gal4 at the RC5 integration site (marked with arrow). (C) Males show a strong enrichment on the X chromosome for MSL3 and MOF-Gal4 stainings (marked with arrow).

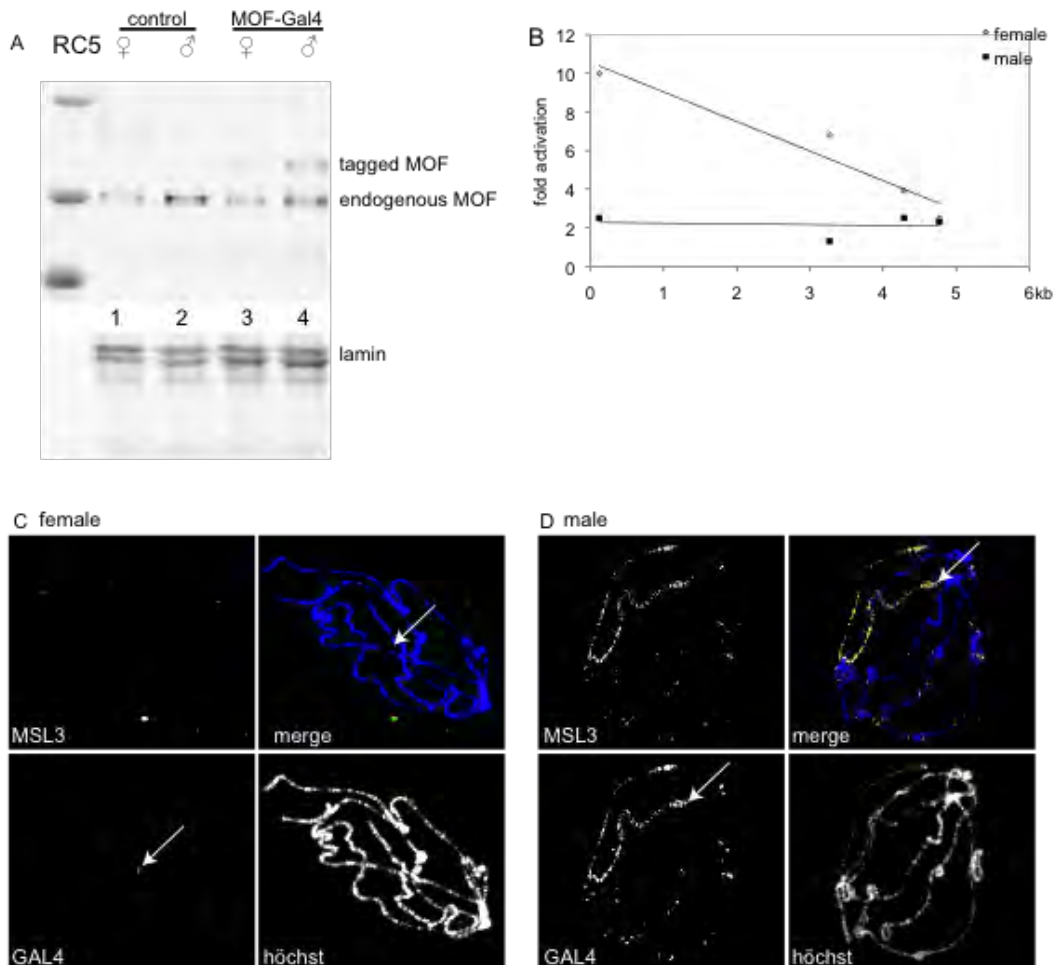


Figure S2: (A-D) Polytene stainings upon MOF recruitment in the RC5 line, zoomed in at the 93B integration site. Polytene squashes and immunofluorescence microscopy was as in figure S1. (A-C) Tethering of MOF-Gal4 to the UAS at 93B (marked with arrows) recruits MSL1, MSL3 and MLE in male salivary glands. (D) In females, MOF-Gal4 is targeted to the RC5 transgene, but no other DCC members (e.g. MSL3) can be detected at this site on polytene chromosomes.

(E,F) The yeast GAL4 activator activates transcription independent of H4K16 acetylation. (F) Expression of the yeast GAL4 activator protein does not induce H4K16ac at the RC5 locus. For reference, the physiological H4K16 acetylation at the native armadillo locus is shown. (G) GAL4 activator protein induces lacZ expression in a nearly sex independent manner. Recruitment of MOF-Gal4 which only contains the DNA binding domain, leads to differential activation in the two sexes. By contrast, recruitment of full-length GAL4 activator leads to similar activation in either sex, independent of H4K16 acetylation.

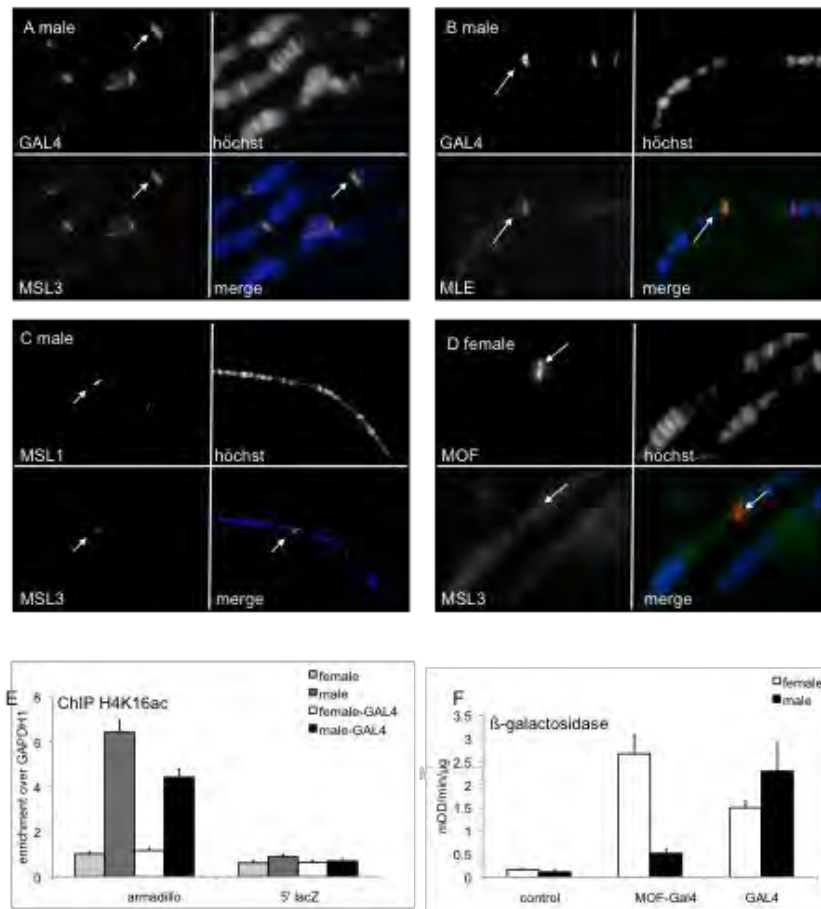


Figure S3: Genome-wide binding of MBD-R2 and MOF in females and males, and MSL1 in males.

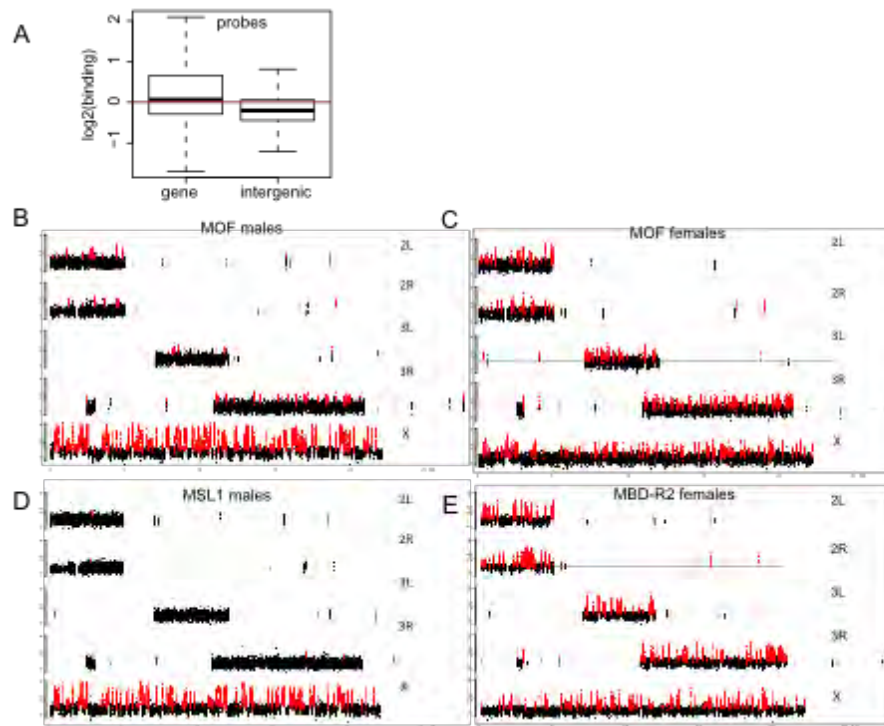
(A) MBD-R2 binds preferentially to genes. ChIP signal (\log_2 IP/input) of oligonucleotide probes mapping to genes or intergenic regions.

(B) Genome-wide binding of MOF in males. Oligonucleotide probe signals are plotted along the major chromosomes (2L, 2R, 3L, 3R, X). Note that our custom-designed tiling array represents most of the X chromosome, but only selected regions of the autosomes (areas of intense signal). Significantly bound probes (tileHMM, Supplementary Methods) are marked in red. Numbers along the X-axis denote the physical position along the chromosomes in megabase pairs (Mb). The y axis denotes the ChIP binding signal as the \log_2 ratio of IP/input.

(C) Genome-wide binding of MOF in females.

(D) Genome-wide binding of MSL1 in males. Note that in contrast to MOF (B) there are only few significant bound probes for MSL1 on the male autosomes.

(E) Genome-wide binding of MBD-R2 in females.



(F-H) MBD-R2 antibody specificity control.

(F) Western blot of SL2 cells treated with dsRNAs against GST or two independent regions of the MBD-R2 gene (M1 and M2) either for 5 or 7 days and probed with α MBD-R2. Lamin was used as loading control. MBD-R2 RNAi knockdown specifically removed two bands stained by the MBD-R2 antibody (presumably, the two bands correspond to the isoform A and B and are not due to degradation).

(G,H) SL2 cells treated with GST or M1 dsRNA for 5 d and stained with α MBD-R2 and α MSL2. Upon MBD-R2 knock-down the signal obtained with the MBD-R2 antibody faded in most of the cells, whereas the MSL2 signal remained.

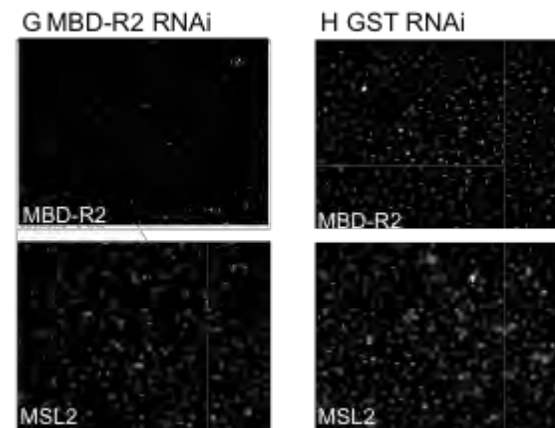
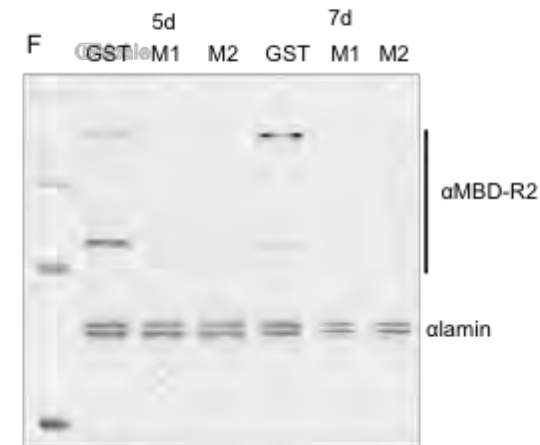


Figure S4: Cumulative profiles of MSL1 binding along genes.

MSL1 binds the gene body and enriches towards the 3' end of genes. Cumulative plot of MSL1. ChIP binding signal along genes were aligned at the transcriptional start site (TSS) and transcriptional termination site (TT). The average binding score is plotted 2 kb bp up- and downstream of the TSS and TT in a sliding window (10 bp step size, 300 bp window). For this analysis, 637 X chromosomal genes were selected, which are minimally 2 kb long, do not overlap with other genes and are expressed. The genes were grouped according to their gene expression level (Chintapalli et al., 2007) into "low" (black), "intermediate" (orange), "high" (blue) and "very high" (purple) categories. The X-axis denotes the distance to the TSS and TT in base pairs.

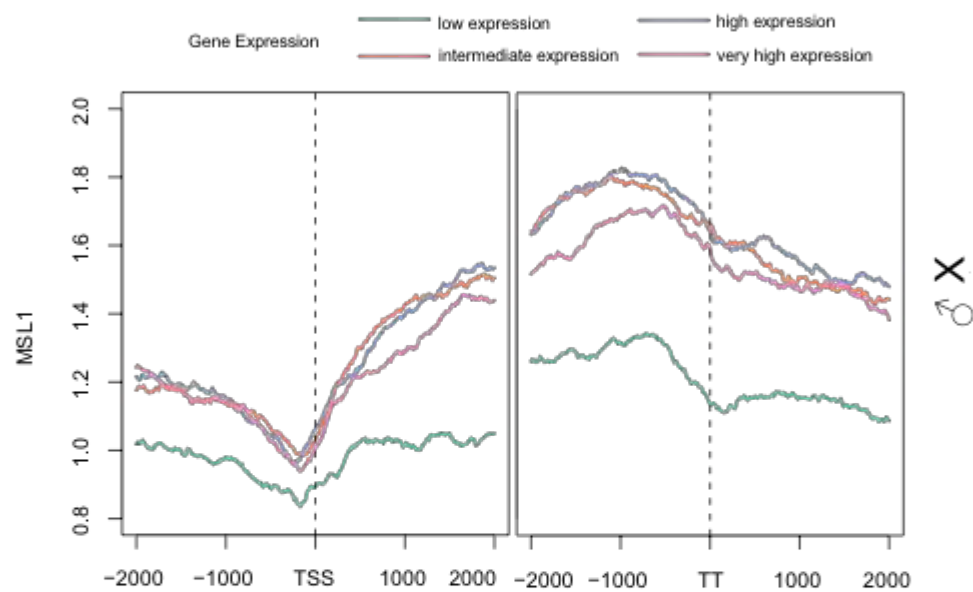


Figure S5: Quality control of biological replicates: Affymetrix transcriptome after MBD-R2 RNAi. Scatter plots of expression change after MBD-R2 RNAi ($\log_2(\text{MBD-R2 RNAi/control RNAi})$) for the two distinct dsRNA amplicons directed against the MBD-R2 transcript (construct 1 and 2). Significantly changed genes (locfdr based on a sam statistic, see Supplementary methods) are marked in red, unchanged genes are marked in blue. Both constructs used to knockdown MBD-R2 caused similar changes in the transcriptome for the two biological replicates. Correlation analyses of each pair of (MBD-R2/control) RNAi transcriptomes of the other replicate arrays performed similar (not shown).

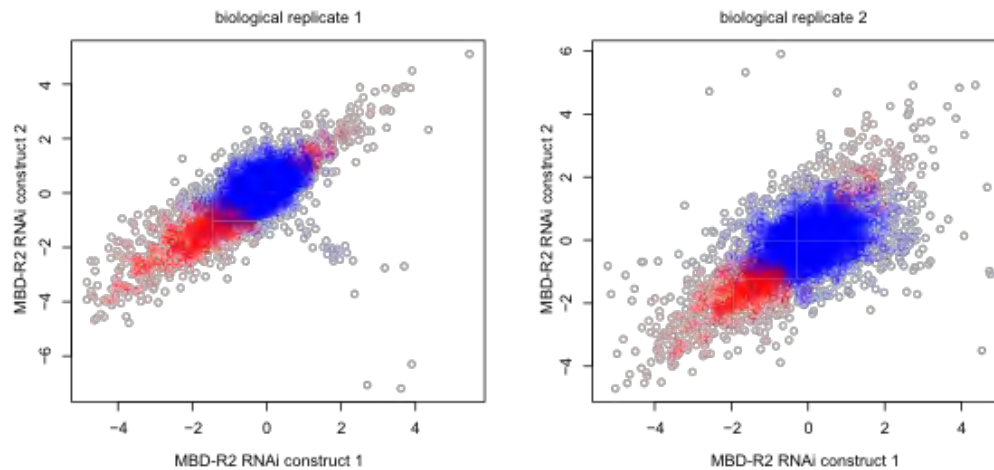
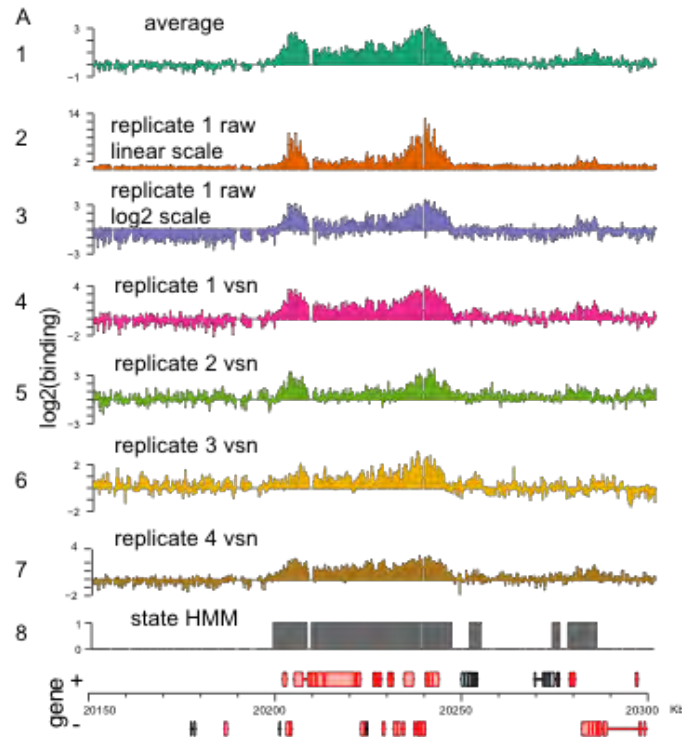
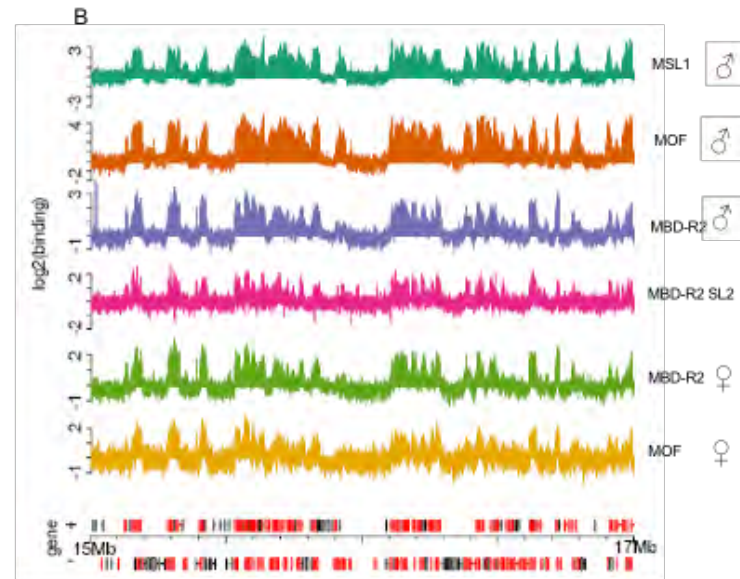


Figure S6: (A) Reproducibility of biological replicates of MBD-R2 ChIP-chip in males along a representative X chromosomal region. Binding profile of 1st biological replicate (lane 2) in raw data linear scale, raw data log2 scale (lane 3), and vsn normalized and log2 scaled (lane 4), vsn normalized log2 scaled of 2nd biological replicate (lane 5), 3rd replicate (lane 6), 4th replicate (lane 7), average of all 4 replicates (lane 1) and tileHMMcalculated "bound" genes (lane 8) over a representative X chromosomal region. The profiles are related to a gene representation at the bottom of the figure, where genes drawn above the line are transcribed from left to right and genes below the line are transcribed from right to left. Active genes are marked in red, inactive genes in black (Chintapalli et al. 2007). Numbers along the X-axis denote the physical position along the chromosomes in kilobase pairs (Kb). The y axis presents the ChIP binding signal as the log2 ratio of IP/input (except in lane 2: linear scale).



(B) Overview of ChIP-chip profiles generated in this study: Binding along a representative region on the X.

Binding of MSL1 (lane 1), MOF (lane 2) and MBD-R2 (lane 3) in adult male flies, MBD-R2 in embryonically-derived SL2 cell line (lane 4), and MBD-R2 (lane 5) and MOF (lane 6) in female adult flies. The profiles are related to a gene representation at the bottom of the figure, where genes drawn above the line are transcribed from left to right and genes below the line are transcribed from right to left. Active genes are marked in red, inactive genes in black (activity state based on Chintapalli et al. 2007). The y axis presents the ChIP binding signal as the log2 ratio of IP/input. Numbers along the X-axis denote the physical position along the chromosomes in megabase pairs (Mb).



Supplemental table

Table S1: Mass spectrometry data for the MOF-MBD-R2 complex purification

protein	scores MOF IP	scores control IP
MOF	2415	0
MBD-R2	700	0
NSL1	572	0
NSL2	142	0
NSL3	438	0
Wds	232	0
dMCRS2	212	0
Rrp4	35	116
Rrp6	118	463
Rrp42	63	142
Dis3	405	723
Z4	61	56
chromator	37	29

Table S2: Cloning primer

cloning primer	Forward	reverse	template clone #
<i>MBD-R2</i>	CCATGGGGCCAT ATGGATACCGCG GAGATCGAAGC	AACAGTTCTAGACTAG TGCTCGCCCACAATGA GGAAACC	SD10773
<i>Mof-FLAG-3'</i>	TGCATGCTCGACT ACAAAGACGACG ACGACAAATAGCA TACGGAACCTGG	TAGGGCCCCCGGGTT TTTCGTTTGGAGGGGT	genomic <i>mof</i>
<i>RsrII-GAL4</i>	ATCGGTCCGAAG CTACTGTCTTCTA TC	TACGGACCGGCCGGC GATACAGTCAA	pAS MOF (Akhtar and Becker, 2000)
<i>KpnI-5xUAS</i>	GAATGGTACCTAT ACTCCGGCGCTC	GAATGGTACCCTCGGA TCCAAGCTT	RC5 construct (in this paper)

Table S3: ChIP-qPCR primer

Locus	Forward	Reverse
<i>armadillo</i> 5'	CACGAACTCCATGTTATTGACTGT C	ATTCTGGGCTGGCATGTAACT
<i>lacZ</i> 5'	GCAACTACTGAAATCTGCCAAG	GTTTTCCCAGTCACGACGTT
<i>lacZ</i> 3'	GCTACATGACATCAACCATATCAG C	GATCCTCTAGAGTCGAGGCC
<i>TL lacZ</i> 3'	TGTTGAAGTGCGAGCGATAC	GGTCGGGATAGTTTTCTTGCG
<i>white</i> 5'	GATCATATCATGATCAAGACATCT AAAGGC	GTGCATCTAGGATCAGCTTAA AATAT
<i>white</i> 3'	TGTGCGTTAGGTCCTGTTTCATTG	CCTGTTCCGAGTGATTAGCGT TAC
<i>gapdh1</i>	GTGACCTACGCAGAAAGCTAG	GCTATTACGACTGCCGCTTTTT C

Table S4: RNAi primer

Construct	Forward	Reverse
<i>MBD-R2_1</i>	TTAATACGACTCACTATAGG GAGATGGAGCCACCAAGT TG	TTAATACGACTCACTATAGGGATGT CGGTCTGGTCATTAGATG
<i>MBD-R2_2</i>	TTAATACGACTCACTATAGG GATC GACGTCGCGT CTGTTG	TTAATACGACTCACTATAGGGA GACAGGATTTGCGCAACTAT
<i>GST</i>	TTAATACGACTCACTATAGG GAGAATGTCCCCTATACTAG GTTA	TTAATACGACTCACTATAGGGAGAA CGCAT CCAGGCACAT TG
<i>msl2_1</i>	TTAATACGACTCACTATAGG GAGAATGGCCCAGACGGCA T AC	TTAATACGACTCACTATAGGGAGAC AGCGATGTGGGCATG TC
<i>msl2_2</i>	TAATACGACTCACTATAGGG TTCCCCTGCTGCCCACAG	TAATACGACTCACTATAGGGCTCTG ACGGGATTGAGGTC

Supplemental experimental procedures

Cloning procedures

An 8 kb *Apal*-*XhoI* fragment containing the intron-free genomic *mof* locus was isolated from clone RP98-11C13 and subcloned into the pBS-SKII. The 147 amino acid GAL4 binding domain (GAL4-DBD) was furnished with *RsrII* restriction sites via PCR and inserted into the unique *RsrII* site 111 bp downstream of the *mof* translation start site. A FLAG tag was added at the 3' end of the *mof* gene. *mof* was expressed from flanking sequences within *KpnI*-*XhoI* fragment and the 3' flanking gene CG3033 was removed via PCR (see primers in supplement) (Fig. 1A). The MOF over-expression construct was created by amplifying the *hsp83* promoter and the 94 amino acid (aa) GAL4-DBD from the *hsp83/GAL4/Dorsal* construct (Flores-Saaib et al., 2001). A *mof* cDNA, containing a 3' FLAG tag, was isolated via PCR and fused to the above described *hsp83*-GAL4-DBD amplicon in a pBluescript yielding *hsp83*-GAL(94)MOF-FLAG. For P-element-mediated transformation the MOF expressing constructs were cloned into the *SmaI* site of the P-element transformation vector pYES (Patton et al., 1992). The RC3 reporter construct was cloned by inserting a 5xUAS of the pUAST-vector (Brand and Perrimon, 1993) into the *KpnI*-*SphI* site. Subsequent digestion with *KpnI* allowed inserting a *KpnI*-*BamHI* *hsp70-lacZ* fragment from the pHZR vector (Gindhart et al., 1995) after blunting.

For antibody production cDNA fragments of MBD-R2 (1-355 aa of the smaller isoform A) and of NSL1 (aa 562-932) were ordered from the *Drosophila* Genomics Resource Center and cloned into the pGEX 2TKN. All primers are listed in the table S2.

The firefly reporter construct pGL3-Tkmod (Gilfillan et al., 2007) was modified as follows. To minimize steric promoter inhibition we introduced 214 bp of the human rDNA locus into the *SmaI* site. For the 5' UAS construct the *KpnI*-5xUAS fragment was amplified via PCR and added into the *KpnI* site (see primers in supplement). The 3' UAS construct was obtained by cloning the same *KpnI*-5xUAS PCR fragment blunt into the *HpaI* site 3' of the firefly luciferase gene. Details are available upon request.

Stable cell lines and protein purification

Kc cells were cotransfected with *phsp83*-GAL(94)MOF-FLAG and pCoHygro (Dignam et al., 1983; Van der Straten et al., 1989). Stable transformants were selected in the presence of 500µg/ml Hygromycin B. A Hygromycin B-resistant clone without *hsp83*-GAL(94)MOF-FLAG expression served as a control. Stable lines were expanded in roller bottles to a final density of 3×10^6 cells/ml.

Nuclear extracts were prepared based on the protocols by Dignam and Roeder (Dignam et al., 1983) and Heberlein and Tjian (Heberlein and Tjian, 1988). PBS-washed Kc cells were resuspended in 5 times of the packed cell volume (PCV) hypotonic buffer A: (10mM HEPES pH 7.6, 15mM KCl, 2mM MgCl₂, 0.1mM EDTA), kept on ice for 15 min and then dounced 10-15 times with a tight B pestle. Buffer B (50mM HEPES pH 7.6, 1M KCl, 30mM MgCl₂, 0.1mM EDTA) was added immediately in a 1:10 ratio to obtain buffer AB. Nuclei were obtained by centrifugation and subsequently resuspended in one PCV of buffer AB. Nuclei were extracted by adding 1/10 volumes of 4M (NH₄)₂SO₄ and incubation for 15 minutes with gentle mixing followed by ultracentrifugation at 48,000 rpm (TLA-55) for 2 hours. The clear supernatant was subsequently precipitated with 0.3g freshly ground (NH₄)₂SO₄ per ml supernatant. Precipitate was resuspended in HEMG150 (25mM HEPES pH7.6, 150mM NaCl, 12.5mM MgCl₂, 0.1mM EDTA, 20% glycerol), dialysed against HEMG150 for 4 hours, snap-frozen and stored at -80°C. All buffers contained 1mM DTT, 1mM Na₂S₂O₅ and 0.5mM PMSF. Final nuclear extracts contained additional protease inhibitors Leupeptin, Pepstatin and Aprotinin.

For Anti-FLAG immunopurification 30µl of anti-FLAG M2 agarose beads (Sigma, A2220) were equilibrated with HEMG150 and mixed with 500µl nuclear extract for 2 hours. Nonspecifically bound material was washed-off three times for 15 minutes with wash buffer (25mM HEPES pH 7.6, 150/250mM NaCl, 12.5mM MgCl₂, 0.1mM EDTA, 0.01% NP-40). Bound proteins were eluted with HEMG150 containing 0.5 mg/mL FLAG peptide and the soluble protein fraction collected after centrifugation.

Mass Spectrometry – Ammoniumbicarbonate (40 mM final concentration) and Trypsin (sequencing grade, modified; Promega; 2 µl of 0.2 µg/µl solution) were added to the eluate and proteins were digested over-night at 37°C while shaking (600 rpm).

For protein identification probes were directly used for nano-ESI-LC-MS/MS. Each sample was first separated on a C18 reversed phase column (75 µm i.d. x 15 cm, packed with C18 PepMap™, 3 µm, 100 Å; LC Packings) via a linear acetonitrile gradient. MS and MS/MS spectra were recorded on an Orbitrap mass spectrometer (Thermo Scientific). The resulting spectra were analyzed via the Mascot™ Software (Matrix Science) using the NCBI nr Protein Database.

β-Galactosidase assay

To measure the expression of *lacZ* a β-galactosidase assay was performed as previously described (Fitzsimons et al., 1999) with modifications. 3-6 adult flies (3-4 days old) were collected and frozen in liquid nitrogen. Flies were grained and

resolved in 50 mM potassium phosphate buffer (pH 7.5) with 1 mM MgCl_2 . The protein concentration was determined according to Bradford. CPRG (Roche Cat No. 10884308001) was used as substrate. The kinetic of this colorimetric assay (mOD/min) was determined with a multi-well plate absorption reader (time frame: minimum 20 min, interval: 20 sec, wavelength: 574 nm; BioTek PowerWave HT) at 37°C and normalized to the protein concentration. Biological replicates were performed.

Luciferase assay

8×10^6 SL2 cells were seeded in 25 cm² flask in 3 ml standard Schneider medium without serum and treated with 40 µg dsRNA (GST, msl2 or MBD-R2). After 1h, 6 ml Schneider medium with serum was added. Cells were incubated for 3 days and subsequently distributed in 6 well plates for transfection (1.5×10^6 per well). A plasmid mix was prepared, containing the activator MOF-Gal4 (MOF promoter), the UAS-containing firefly luciferase reporter and the Renilla luciferase reporter (both tk promoter) for normalisation. Transfection was performed according to the Effectene transfection kit (Qiagen, #301427). After 2.5 days cells were harvested, lysed by freezing and thawing in 300 mM KCl, 50 mM HEPES pH 7.6, 0.5 mM EDTA, 0.1% NP40 and 1 mM DTT in the presence of protease inhibitors. 1/10 of the lysate was assayed for luciferase activity in a Lumat LB 9501 (Berthold), according to the Promega Dual-Luciferase assay Manual (#E1960).

Chromatin preparation and chromatin immunoprecipitation

Chromatin was prepared from sex-sorted adult flies. The protocol from Negre et al. (Negre et al., 2006) was applied with some modification: 300 mg of 3-4 days old flies were crushed in 5 ml A1 (60 mM KCl, 15 mM NaCl, 4 mM MgCl_2 , 15 mM HEPES (pH 7.6), 0.5% Triton X-100, 0.5 mM DTT, 10 mM sodium butyrate, protease inhibitor cocktail (Roche, Cat No. 04693132011)) with a final concentration of 2.35% formaldehyde at 18°C for 15 min. Fixation was stopped with a final concentration of 250 mM glycine. After 3 washes with A1 the pellet was equilibrated in lysis buffer (140 mM NaCl, 15 mM HEPES (pH 7.6), 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.5% sodium deoxycholate, protease inhibitor cocktail, 0.5mM DTT). The pellet was suspended in lysis buffer with 0.1% SDS and 0.5% N-lauroylsarcosine and rotated for 10 min at 4°C. Chromatin was sonicated (Branson (microtip), 4 times at 21 watt) in the presence of glass beads (212-300 µm) in a volume of 2 ml in 15 ml falcon tubes. Cell debris was removed by centrifugation and the chromatin containing supernatant was stored in aliquots at -80°C.

For chromatin immunoprecipitation (ChIP) the DNA concentration of the chromatin was determined and 7.5 µg of DNA was used per IP in 500 µl lysis buffer plus 250 µl PBS (1xPBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, protease inhibitor cocktail). Chromatin was pre-cleared with protein-A and -G beads at 4°C for 30 min. After pre-clearing an aliquot was saved for input DNA purification. Immunoprecipitation was carried out at 4°C overnight. The amount of antibody used per ChIP was adjusted individually. For the pull down 30 µl protein A and G beads were used per ChIP at 4°C for 3 h. Beads were washed 5 times in PBS, rinsed once in EB buffer (10 mM Tris-Cl, pH 8.5). De-crosslinking was performed in 100 µl EB with 0.6% SDS and a final concentration of 0.5 µg proteinase K and RNase A. Digest was carried out at 55°C for 3 h and de-crosslinking at 65°C for 6 h. ChIP and input DNA were purified with phenol/chloroform using MaXtract High Density tubes (Qiagen, Cat No. 129056) and subsequent ethanol precipitation.

Chromatin preparation of SL2 cells was performed as previously described in Straub et al. (2008).

ChIP-chip data analysis

ChIP-chip data analysis was essentially performed as in Straub et al. 2008. Briefly, input and IP DNA were amplified using the WGA kit (Sigma) according to an online protocol (<http://www.epigenome-noe.net/researchtools/protocol.php?protid=30>). Labeling and hybridization to NimbleGen arrays was carried out at ImaGenes (Berlin, Germany). We used a custom array layout (approx. 1 probe/100bases, isothermal selection) comprising the euchromatic part of the entire X chromosome, 5 Mb of 2L, 2R and 3L, respectively, as well as 10 Mb of 3R. Data analysis was performed using R/Bioconductor (www.Rproject.org; www.bioconductor.org). Raw signals of corresponding biological replicates were normalized and log2 transformed using the „vsn” package (Huber et al., 2002). Averaged binding scores per gene present the enrichments (log2 ratio of IP and input) normalized on the gene length. Enrichment statistics (IP versus input signals) were computed using the „sam” algorithm within Bioconductor (Tusher et al., 2001). Fdr values of the sam statistic were determined using „locfdr” (Efron, 2007). Region summarization was performed using the HMM algorithm of tileHMM (Humburg et al., 2008) with the following parameters: fragment size of 700, maximal gap of 400, minimal length of 400 and minimal score of 0.8. Genes were considered „bound” significantly with more than 5 tileHMM „bound” probes. Hierarchical cluster analysis of the binding pattern across genes was carried out using the „hclust” package of R. p values for figure 7B were Benjamini-Hochberg

adjusted. All data correspond to *Drosophila* genome version dm2 and annotation version gadfly 4.3.

Transcriptome analysis

Cultivation of the male *Drosophila* cell line SL2 and RNA interference (RNAi) of target genes were carried out as described previously (Straub et al., 2005). In brief, 1.5×10^6 SL2 cells were incubated with 10 μ g dsRNA for 1 hour in serum-free medium. Sequences of primers used for dsRNA production are listed in the table S4. After addition of serum-containing medium cells were incubated for 7 days at 26°C before RNA extraction. Preparation of chromatin extracts and western blot confirmation of target gene knockdown has been described previously. Depletion efficiency was quantified using a Li-Cor Odyssey system with α -tubulin as reference.

RNA was isolated using Trizol (Invitrogen) followed by a purification using RNeasy kit (Quiagen) according to the instructions of the suppliers. RNA labeling and cDNA hybridization to a *Drosophila* Genome GeneChip 2.0 was performed by ImaGenes (Berlin, Germany). Two biological replicate experiments were performed with a total of four control RNAi (3x GST RNAi and 1x mock RNAi) and six MBD-R2 RNAi (3x with each dsRNA construct). Data analysis was performed using R/Bioconductor (www.Rproject.org; www.bioconductor.org). Intensity values were normalized, summarized and log2 transformed using the „gcrma“ package (Wu and Irizarry, 2005). Other normalization methods (vsN, quantile) were also tested and performed similarly. Quality control assessment was performed using the R package “arrayQualityMetrics” (Kauffmann et al., 2009). We did not observe significant batch effects of the biological replicates. The transcriptome changes are similar upon treatment of two distinct dsRNA constructs targeting the MBD-R2 transcripts (Figure S3). Thus, the expression values for the control samples and the MBD-R2 RNAi samples were averaged, respectively. Significant change of gene expression was calculated using locfdr based on a sam statistics (Efron, 2007; Tusher et al., 2001). Genes were considered “differentially expressed” with an fdr cutoff of 0.35. Alternatively, an eBayes moderated t test or a limma statistic followed by multiple testing correction (locfdr) gave similar results. The results are robust to various parameters in data analysis, as assessed by choosing varying thresholds. The expression data of adult flies were taken from Chintapalli et al. and analyzed using the same algorithms of the MBD-R2 data set. All values of the expression set data are log2 normalized with a theoretical dynamic range of $2^{\exp 16}$ (Affymetrix.com).

Supplemental references

- Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401-415.
- Dignam, J.D., Martin, P.L., Shastry, B.S., and Roeder, R.G. (1983). Eukaryotic gene transcription with purified components. *Methods Enzymol* 101, 582-598.
- Efron (2007). Correlation and large-scale simultaneous significance testing. *J. Amer. Statist. Assoc* 102, 93-103.
- Fitzsimons, H.L., Henry, R.A., and Scott, M.J. (1999). Development of an insulated reporter system to search for cis-acting DNA sequences required for dosage compensation in *Drosophila*. *Genetica* 105, 215-226.
- Flores-Saaib, R.D., Jia, S., and Courey, A.J. (2001). Activation and repression by the C-terminal domain of Dorsal. *Development* 128, 1869-1879.
- Gindhart, J.G., Jr., King, A.N., and Kaufman, T.C. (1995). Characterization of the cis-regulatory region of the *Drosophila* homeotic gene *Sex combs reduced*. *Genetics* 139, 781-795.
- Heberlein, U., and Tjian, R. (1988). Temporal pattern of alcohol dehydrogenase gene transcription reproduced by *Drosophila* stage-specific embryonic extracts. *nature* 331, 410-415.
- Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 18 Suppl 1, S96-104.
- Humburg, P., Bulger, D., and Stone, G. (2008). Parameter estimation for robust HMM analysis of ChIP-chip data. *BMC Bioinformatics* 9, 343.
- Kauffmann, A., Gentleman, R., and Huber, W. (2009). arrayQualityMetrics--a bioconductor package for quality assessment of microarray data. *Bioinformatics* 25, 415-416.
- Negre, N., Hennetin, J., Sun, L.V., Lavrov, S., Bellis, M., White, K.P., and Cavalli, G. (2006). Chromosomal distribution of PcG proteins during *Drosophila* development. *PLoS Biol* 4, e170.
- Patton, J.S., Gomes, X.V., and Geyer, P.K. (1992). Position-independent germline transformation in *Drosophila* using a cuticle pigmentation gene as a selectable marker. *Nucleic Acids Res* 20, 5859-5860.
- Straub, T., Gilfillan, G.D., Maier, V.K., and Becker, P.B. (2005). The *Drosophila* MSL complex activates the transcription of target genes. *Genes Dev* 19, 2284-2288.
- Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-5121.

Van der Straten, A., Johansen, H., Rosenberg, M., and Sweet, R. (1989). Introduction and constitutive expression of gene products in cultured *Drosophila* cells using hygromycin B selection. *Methods in Molecular and Cellular Biology* 1.

Wu, Z., and Irizarry, R.A. (2005). Stochastic models inspired by hybridization theory for short oligonucleotide arrays. *J Comput Biol* 12, 882-893.

3.2 Dosage compensation and the global re-balancing of aneuploid genomes (Review)

Matthias Prestel, Christian Feller and Peter B. Becker

3.2.1 Summary and own contribution

In this review written by Dr. Matthias Prestel (50%), Prof. Peter Becker (25%) and myself (25%), we discuss the implications of the findings presented in section 3.1 in the context of general strategies to balance aneuploid genomes. It has been proposed before that three types of compensatory mechanisms are in place to respond to genome aneuploidies. According to this model, dosage compensation in male flies is a composite of a more general feedback or buffering principle and an additional feed-forward mechanism exerted by the MSL-DCC. We expand this model by suggesting that the precise two-fold activation during dosage compensation is achieved by a balancing of counteracting activating and repressive activities.

In addition to contributing to the text, I prepared the figures 1, 3 and 4.

3.2.2 Published review article

REVIEW

Dosage compensation and the global re-balancing of aneuploid genomes

Matthias Prestel, Christian Feller and Peter B Becker*

Abstract

Diploid genomes are exquisitely balanced systems of gene expression. The dosage-compensation systems that evolved along with monosomic sex chromosomes exemplify the intricacies of compensating for differences in gene copy number by transcriptional regulation.

Complex genomes are more than just the sum of their genes, but are rather complex regulatory systems in which the expression of each individual gene is a function of the activity of many other genes, so that the levels of their protein products are maintained within a narrow range. Such homeostasis favors the maintenance of the appropriate stoichiometry of subunits in multiprotein complexes or of components in signal transduction pathways, and defines the 'ground state' of a cell [1]. In diploid genomes, both alleles of a gene are usually active and this 'double dose' of each gene is figured into the equation. Thus, deviations from diploidy, such as the deletion or duplication of genes or of larger chromosomal fragments (aneuploidy), unbalance the finely tuned expression of the genome. Segmental aneuploidies of this kind can arise from failed or faulty repair of chromosomal damage due to irradiation, chemical insult or perturbation of replication, or from illegitimate recombination during meiosis. Loss or duplication of entire chromosomes (monosomy or trisomy, respectively) can arise from non-disjunction during cell division. Depending on the extent of the aneuploidy and on the genes affected, the fine balance of *trans*-acting factors and their chromosomal binding sites that define the gene-expression system is disturbed, and the fitness of the cell or organism challenged.

Often, aneuploidies have been associated with a variety of developmental defects and malignant aberrations,

such as Down syndrome or certain breast cancers (reviewed in [2,3]). The phenotypes associated with changes in gene copy number can not only be the result of the deregulation of the affected gene(s), but may also reflect *trans*-acting effects on other chromosomal loci or even more global alterations of the entire regulatory system. This is particularly true if genes coding for regulatory factors, such as transcription factors, are affected (reviewed in [4,5]).

Strategies for re-balancing aneuploid genomes

Genome-wide studies in different organisms reveal that the expression of a substantial number of genes directly correlates with gene dose (the primary dosage effect) [6]. In other cases, the measured expression levels do not reflect the actual copy number, as compensatory mechanisms aimed at re-establishing homeostasis take effect [4,5]. Imbalances due to aneuploidy may be compensated for at any step of gene expression from transcription to protein stability. Excess subunits of multiprotein complexes that are not stabilized by appropriate interactions are susceptible to degradation (see [1] for a discussion of compensation at the protein level). Dosage-compensation mechanisms at the level of transcription are versatile, intricate, and in no instance are they fully understood.

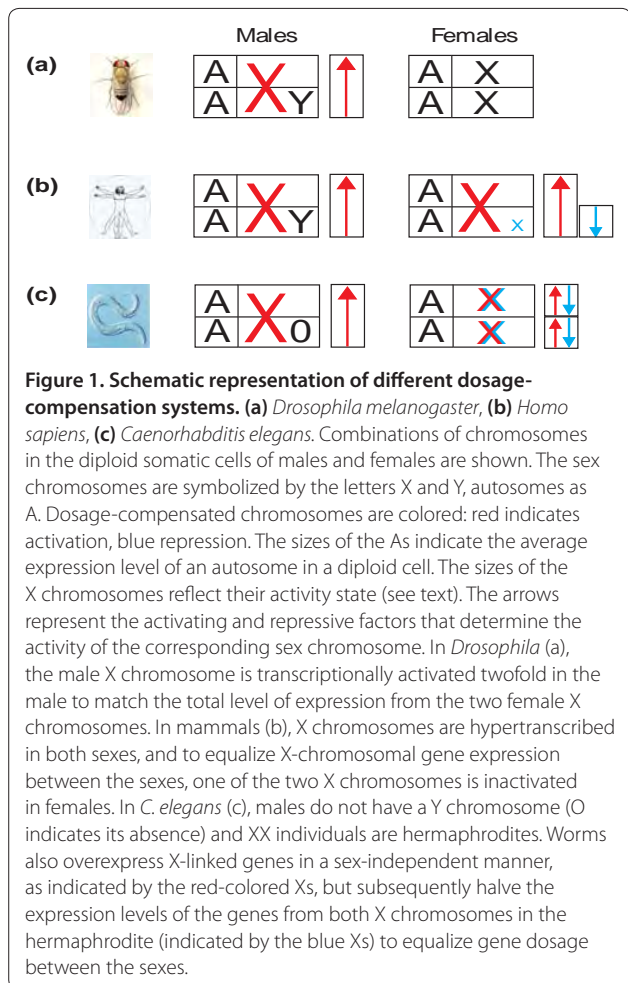
In principle, three types of compensatory responses to aneuploidies are recognized: buffering, feedback, and feed-forward, which may act individually or, more likely, in combination [7]. Oliver and colleagues [7] define buffering as 'the passive absorption of gene dose perturbations by inherent system properties'. Currently, the nature of this general or 'autosomal' buffering is unknown, but its existence can be deduced from comparing gene expression to DNA copy number in healthy and aneuploid genomes [8-11]. The system properties referred to by Oliver and colleagues can be considered as the sum of the biochemical equilibria of the system 'living cell', which are predicted to moderate the effect of the reduction of one component. Apparently, the deletion of one gene copy (that is, a twofold reduction in gene expression) can be partially compensated for by increasing the steady-state mRNA levels originating from

*Correspondence: pbecker@med.uni-muenchen.de
Adolf-Butenandt-Institute and Centre for Integrated Protein Science (CiPSM),
Ludwig-Maximilians-University, Schillerstrasse 44, 80336 Munich, Germany

the remaining allele by, on average, 1.5-fold [7,11]. Interestingly, Stenberg and colleagues [11] observed that buffering appears to compensate for deficiencies better than for gene duplications, which leaves open the existence of a general sensor of monosomy that mediates the effect. A general buffering will also ameliorate the consequences of widespread mono-allelic gene expression due to parental imprinting (cases where a single allele is expressed, depending on whether it is inherited from the father or mother) [12].

In contrast to the general and nonspecific buffering just described, a 'feedback' mechanism would be defined as gene-specific - sensing and readjusting the levels of specific molecules by appropriate, specific mechanisms. Finally, 'feed-forward' anticipates the deviation from the norm and hence can only be at work in very special circumstances. Prominent examples where feed-forward scenarios are applicable are the widely occurring monosomies in the sex chromosomes of heterogametic organisms (for example, the XX/XY sex-chromosome system), which are present in each and every cell of the species.

In contrast to aneuploidies that arise spontaneously, these 'natural' monosomies and their associated dosage-compensation mechanisms are the products of evolution. Research on dosage-compensation mechanisms associated with sex chromosomes continues to uncover unexpected complexities and intricacies. The somatic cells of the two sexes of the main model organisms of current research - mammals, nematode worms (*Caenorhabditis elegans*) and fruit flies (*Drosophila melanogaster*) - differ in that those of females are characterized by two X chromosomes, while those of males have one X and one Y chromosome (mammals and *Drosophila*); or one sex (XX) is a hermaphrodite and the males have just a single X and no Y chromosome (X0) (*C. elegans*) [13]. Remarkably, different dosage-compensation strategies for balancing gene expression from the X chromosome between the sexes have evolved independently in these three cases (Figure 1), as we shall discuss in this article. There is increasing evidence that in all three cases, the transcription of most genes on the single male X chromosome is increased roughly twofold [14-16]. In fruit flies, this upregulation of the X chromosome is limited to males. In mammals and worms, however, the X chromosomes appear to be also upregulated in the XX sex, which necessitates additional compensatory measures. In female mammals, one of the X chromosomes is globally silenced, whereas in hermaphrodite worms, gene expression on both X chromosomes is downregulated by about 50% (Figure 1). An emerging principle is that the net fold-changes of dosage compensation are not achieved by a single mechanism (that is, there is no simple switch for 'twofold up'), but by integration of activating and repressive cues, as discussed later.



In what follows we summarize recent insight into the dosage-compensation mechanisms of the XX/XY sex chromosome systems, which nicely illustrate the evolution of global, genome-wide regulatory strategies. However, compensation systems of this type are not absolutely required for the evolution of heterogametic sex. Birds, some reptiles, and some other species use the ZW/ZZ sex-chromosome system, which does not use the mechanism of chromosome-wide transcriptional regulation to compensate for monosomy [17-19].

Dosage compensation of sex chromosomes reveals the balancing capacity of chromatin

The sex chromosomes of the XX/XY system are thought to have originated from two identical chromosomes in a slow process that was initiated by the appearance of a male-determining gene. In order to be effective, this gene should be propagated only in males, which was achieved by evolving a Y chromosome that was specifically propagated through the male germline. The necessary suppression of recombination between this 'neo-Y' chromosome

and the corresponding sister chromosome (which would become the future 'neo-X') favored the accumulation of mutations, deletions and transposon insertions, an erosive process that led to loss or severe degeneration of Y chromosomes [20-24]. The progressive erosion of the evolving Y left many X-chromosomal genes without a corresponding copy on the Y chromosome (the hemizygous state). The initial consequences of gene loss on the Y chromosome may have been absorbed by the intrinsic biochemical buffering properties of the cell noted above [11]. However, when the majority of genes on the X chromosome lost their homologs on the Y chromosome the co-evolution of regulatory processes to overcome the reduced gene dose - that is, dosage-compensation systems - increased the fitness of the organisms. These dosage-compensation systems are likely to originate in the male sex (XY or XO in the examples discussed here), as it is in males that factors acting in a dose-dependent manner (such as transcription factors, chromatin constituents and components of signal-transduction cascades) would become limiting [25,26].

A logical adaptation to ensure the survival of males would be the increased expression of X-chromosomal genes [6]. This intuitively obvious mechanism has long been known in *Drosophila*. Observing the specialized polytene chromosomes in larvae (which are composed of thousands of synapsed chromatids arising from repeated DNA replication without chromosome segregation), Mukherjee and Beermann [27] were able to directly visualize nascent RNA and found that the single X chromosome in males gave rise to almost as much RNA as two autosomes. Recent genome-wide expression analyses confirmed these early observations [28,29] and further genome-wide studies suggest that this mechanism may also operate in *C. elegans* and mammals [14-16]. For these species neither the mechanism of this chromosome-wide regulation nor the factors involved are known.

For *Drosophila*, however, thanks to decades of outstanding genetics exploring male-specific lethality, we know at least a few of the prominent players. Here, the twofold stimulation of transcription on the X chromosome is mediated by the male-specific assembly of a dosage-compensation complex (the Male-Specific-Lethal (MSL) complex), a ribonucleoprotein complex that associates almost exclusively with the X chromosome (reviewed in [30]; Figure 2). Most subunits of the MSL complex are found in both sexes of *Drosophila*, except for the key protein MSL2 and the noncoding *roX* (RNA-on-the-X) RNAs, which are only expressed in males (Figure 2), thus leading to the assembly of the MSL complex exclusively in male cells. The MSL complex associates with the transcribed regions of target genes in a multi-step process that has been reviewed elsewhere [31-33]. Key to the stimulation of transcription is the

MSL-complex subunit MOF (Males-absent-on-the-first; also known as KAT8, lysine acetyltransferase 8), a histone acetyltransferase with specificity for lysine 16 in the amino-terminal tail of histone 4 (H4K16ac). Acetylation of this residue is known to reduce interactions between nearby nucleosomes and leads to unfolding of nucleosomal fibers *in vitro* [34,35].

Whereas the action of the dosage-compensation complex in *Drosophila* is limited to males, in *C. elegans* and mammals the unknown factors that stimulate X-chromosomal transcription appear to be active in the hermaphrodite and the female, as well as in males. If, however, X activation re-balances the male genome in these species, it follows that in the XX sex, having two hyperactive X chromosomes relative to the autosomes must be suboptimal [36]. Consequently, further compensation is needed. Mammals have evolved a strategy of inactivating one of the female X chromosomes to achieve a level of X-chromosome gene expression closely resembling that from the single X in males (reviewed in [37]; Figure 1b). Which X is inactivated is random, and inactivation starts with the stable transcription of the long, non-coding *Xist* (Xi-specific transcript) and *RepA* (repeat A) RNAs from a complex genetic region on the future inactive X (Xi) called the X-inactivation center. Subsequently, *Xist* RNA - possibly in complex with undefined protein components - spreads to coat the entire Xi. Silencing involves the recruitment and action of the Polycomb silencing machinery via the *Xist* and *RepA* RNAs [38,39], followed by reinforcement through the incorporation of histone variants, removal of activating histone modifications and DNA methylation [37]. Remarkably, the independent evolution of nematode worms arrived at a very different solution to the problem. *C. elegans* equalizes the gene dose by halving the expression levels of genes on both X chromosomes in the hermaphrodite, using a large dosage-compensation complex containing components of the meiotic/mitotic condensin. The involvement of condensins may point to regulation at the level of chromatin fiber compaction ([40] and references therein). The scenario shown schematically in Figure 1c for *C. elegans* suggests that dosage compensation in this species involves a twofold increase in X-linked transcription in both sexes, which is opposed by a twofold repression in hermaphrodites. The underlying mechanisms are still mysterious.

This short summary of the three very different dosage-compensation systems reveals two common denominators. First, they all adapt factors and mechanisms, which are already involved in other regulatory processes, for the compensation task by harnessing them in a new molecular context. Furthermore, these factors are all known for their roles in modulating chromatin structure. It seems that chromatin can adopt a variety of structures

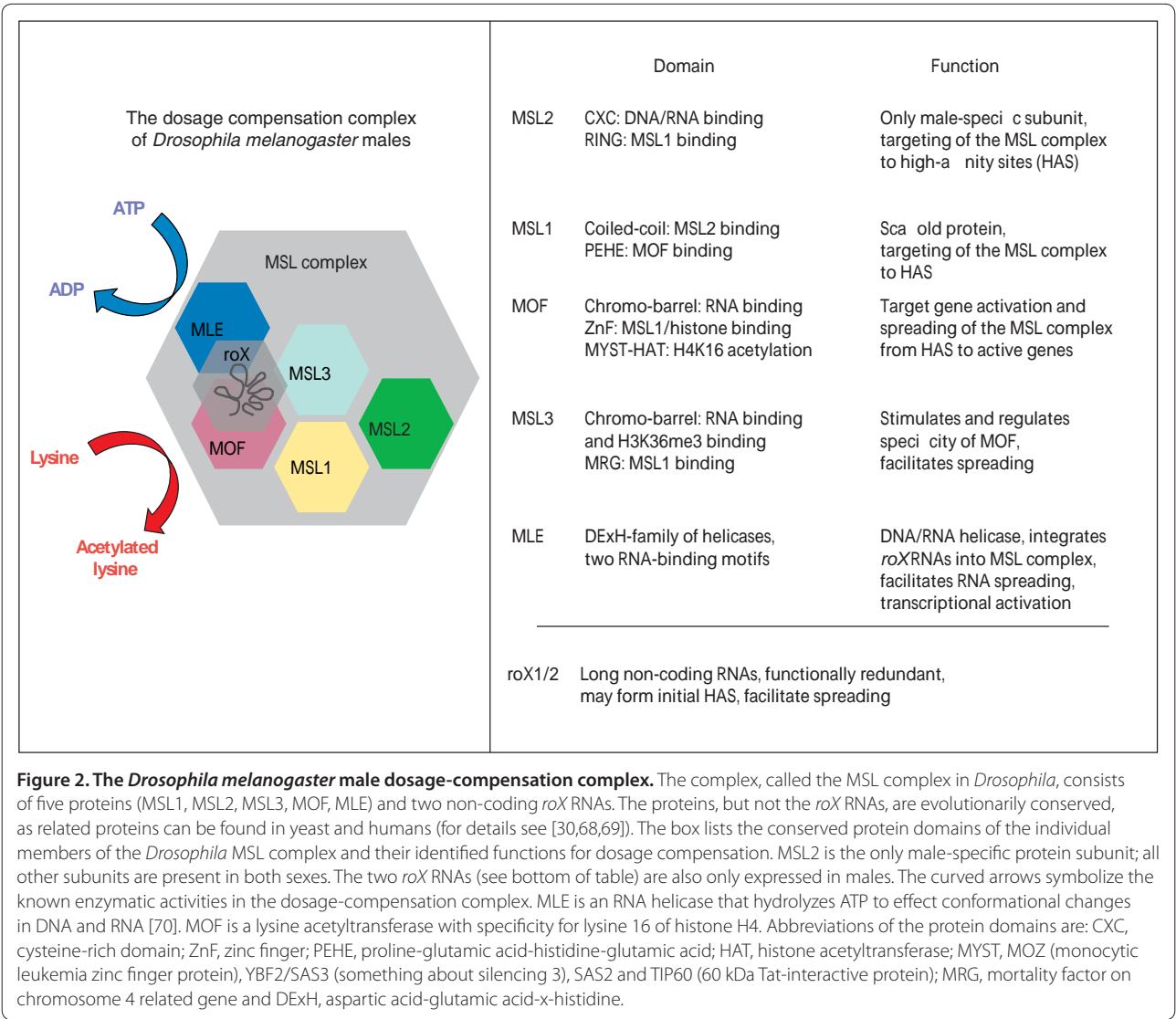


Figure 2. The *Drosophila melanogaster* male dosage-compensation complex. The complex, called the MSL complex in *Drosophila*, consists of five proteins (MSL1, MSL2, MSL3, MOF, MLE) and two non-coding roX RNAs. The proteins, but not the roX RNAs, are evolutionarily conserved, as related proteins can be found in yeast and humans (for details see [30,68,69]). The box lists the conserved protein domains of the individual members of the *Drosophila* MSL complex and their identified functions for dosage compensation. MSL2 is the only male-specific protein subunit; all other subunits are present in both sexes. The two roX RNAs (see bottom of table) are also only expressed in males. The curved arrows symbolize the known enzymatic activities in the dosage-compensation complex. MLE is an RNA helicase that hydrolyzes ATP to effect conformational changes in DNA and RNA [70]. MOF is a lysine acetyltransferase with specificity for lysine 16 of histone H4. Abbreviations of the protein domains are: CXC, cysteine-rich domain; ZnF, zinc finger; PEHE, proline-glutamic acid-histidine-glutamic acid; HAT, histone acetyltransferase; MYST, MOZ (monocytic leukemia zinc finger protein), YBF2/SAS3 (something about silencing 3), SAS2 and TIP60 (60 kDa Tat-interactive protein); MRG, mortality factor on chromosome 4 related gene and DEXH, aspartic acid-glutamic acid-x-histidine.

with graded activity states, which can be used either to completely switch off large chromosomal domains or to fine-tune transcription (either up or down) in the twofold range. Dosage compensation therefore integrates with other aspects of chromatin organization. In *Drosophila*, the male X chromosome that accumulates the H4K16 acetylation mark is particularly sensitive to mutations in general chromatin organizers. Prominent among these is the zinc finger protein Su(var)3-7 (suppressor of variegation 3-7), a heterochromatin constituent known to bind HP1 (heterochromatin protein 1). Normal levels of Su(var)3-7 are required for proper dosage compensation and to ensure the selective binding of the dosage-compensation complex to the X chromosome [41-43]. The male X polytene chromosome bloats when Su(var)3-7 levels are reduced and condenses when the protein is in excess. These changes in chromatin

condensation depend on a functional dosage compensation complex, suggesting that the MOF-catalyzed acetylation of histone 4, and subsequent unfolding effect of H4K16ac, is constrained by as yet unknown counter-acting factors (Figure 3a), conceivably by ones that promote chromatin compaction. Selective, massive unfolding of the dosage-compensated male X chromosome in *Drosophila* is also observed when the nucleosome remodeling factor (NURF) is inactivated [44,45]. Nucleosome remodeling by NURF may thus also serve to counteract excessive unfolding due to H4K16 acetylation. Tamkun and colleagues [46] suggested that NURF might achieve this task by maintaining sufficiently high histone H1 levels on the X chromosome. Clearly, the degree of chromatin compaction can be adjusted by integration of unfolding and compacting factors.

Harnessing MOF for dosage compensation

Further analysis of the role of *Drosophila* MOF in dosage compensation suggests that it may affect gene expression by modulating the productivity of the transcription machinery in the chromatin context. Although MOF is able to acetylate non-histone substrates [47,48], its main substrate in the context of dosage compensation is the strategic H4K16. Biochemical studies showed that this modification interferes directly with the folding of the nucleosomal chain into 30-nm fibers *in vitro* [35,49]. Accordingly, H4K16 acetylation by MOF has the potential to counteract chromatin-mediated transcriptional repression [50,51] (Figure 3a). In the simplest scenario, the only task of the MSL complex in *Drosophila* would be to enrich MOF on the X chromosome relative to the autosomes. However, studies of the effect of MOF in yeast or in a cell-free chromatin transcription system showed that H4K16 acetylation does not automatically increase transcription by twofold, but by many-fold [50]. This strong activation potential of MOF can also be visualized in *Drosophila*. We recently established *Drosophila* lines in which MOF is tethered to a β -galactosidase reporter gene engineered to reside on an autosome [51]. Sorting adult flies according to sex allowed comparison of MOF-dependent reporter gene stimulation in male flies, where MOF is part of the dosage-compensation complex, and in females, where its molecular context was initially unknown. In females, MOF recruitment stimulated transcription from a proximal promoter by an order of magnitude. The effect faded with increasing distance between recruitment site and transcription start site and therefore appears to be related to local chromatin opening by promoter-bound co-activators.

By contrast, the molecular context of the MSL complex in males restricted the activation effect of MOF to the twofold range reminiscent of dosage compensation, and this effect was observable over a distance of 5 kb [51]. Notably, similar H4K16 acetylation levels accompanied the very different activation modes in the two sexes. So it seems that the activation potential of H4K16 acetylation revealed in females is constrained in males. Ectopic assembly of the MSL complex in females by expression of MSL2 constrained the strong activation to a twofold range [51]. We concluded from these and further studies that the *Drosophila* dosage-compensation complex achieves a twofold activation of transcription by integrating activating and repressive principles [51].

MOF serves as an example of the principle that dosage compensation employs chromatin modifiers that are also functional in other contexts. MOF is expressed at only slightly lower levels in females than in males, and it also resides in at least one other complex in addition to the MSL complex. Mendjan *et al.* [52] first reported the existence of an alternative complex (the NSL complex,

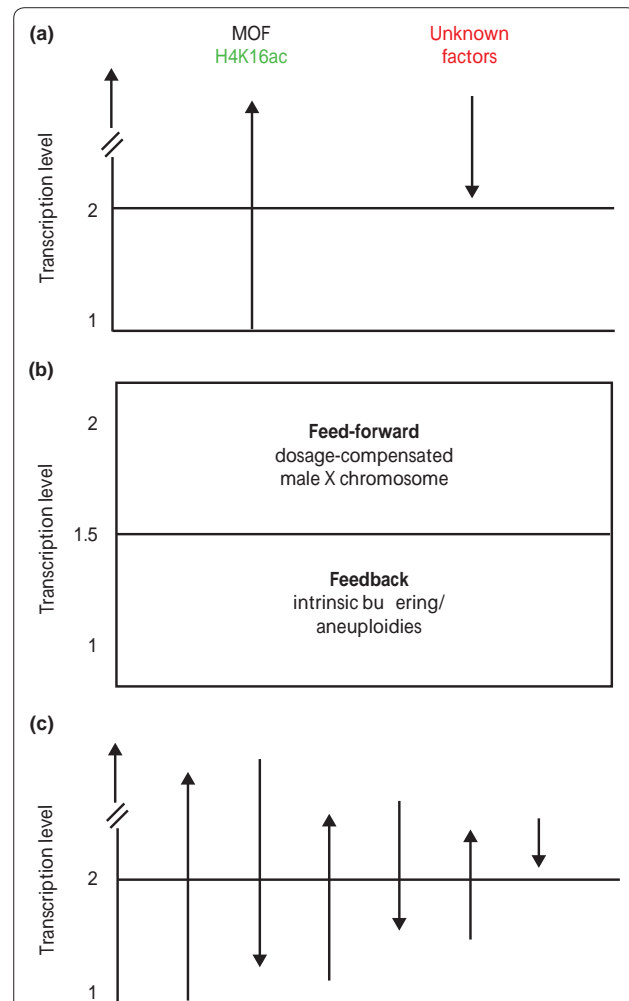


Figure 3. Possible mechanisms for dosage compensation.

(a) The twofold activation of the single male X chromosome in *Drosophila* could be achieved by a large, MOF-dependent activation of transcription through H4K16 acetylation and its counteraction by yet unknown factors, mediated by the dosage-compensation complex in males [51]. In (a,b), transcriptional level 1 refers to the normal regulated level of transcription from a single uncompensated X chromosome in females. (b) Furthermore, the twofold activation of the male X chromosome could be achieved by a combination of mechanisms: a general buffering/feedback component and a dedicated feed-forward mechanism (dosage compensation as suggested in (a)) [7]. The effects of these two processes could add up to the expected twofold compensation required to equalize the expression of X-linked genes between the sexes. (c) Precise transcription levels could result from negotiation between a number of activating and repressive factors (up and down arrows). In this instance, transcriptional level 1 refers to a 'basal' transcription state. This hypothetical model assumes that additional factors beyond those mentioned in (a) and (b) contribute to final transcription levels, such as male-enriched protein kinases, heterochromatin components, chromatin remodelers, and others (for details, see text).

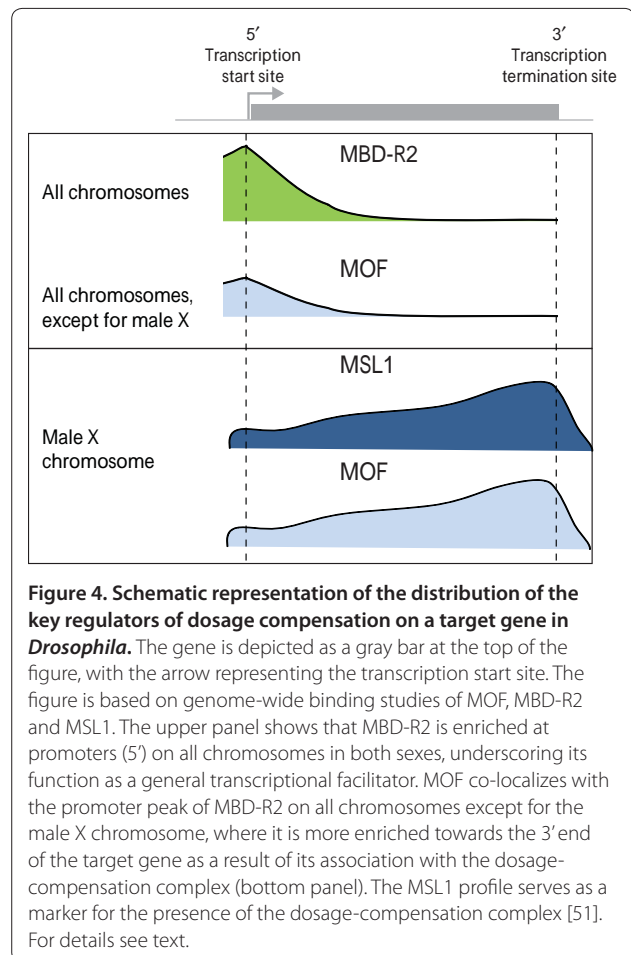
for 'Non-Specific-Lethal') in mixed-sex embryos and male cells of *Drosophila*, which contained a number of poorly characterized nuclear proteins and two

components of nuclear pores [52]. The closely related MOF-MBD-R2 complex, purified by us from female *Drosophila* cells [51], shares several prominent components with the NSL complex, including WDS (Will Die Slowly, a homolog of mammalian WDR5 (WD repeat-containing protein 5), dMCRS2 (microspherule protein 1), a forkhead-associated domain protein, and MBD-R2 (an uncharacterized protein harboring similarity to methyl-CpG-binding domains) [53]. In contrast to the NSL complex, the MOF-MBD-R2 complex does not contain nuclear pore components [51].

The evidence so far suggests that the MOF-MBD-R2 complex provides the molecular context for the strong activation elicited by MOF in females. Globally, MOF co-localizes with MBD-R2 to active genes with enrichment towards their 5' ends on all chromosomes in male and females, except for the male X chromosome (Figure 4). In male *Drosophila* cells, MOF is enriched on the X chromosome, where it co-localizes with MSL-complex components (such as MSL1) with a bias towards the 3' end (Figure 4). In male *Drosophila* cells, MOF apparently distributes dynamically between the two complexes. Ectopic expression of MSL2 in female cells, which leads to assembly of a dosage-compensation complex, re-localizes MOF from the autosomes to the X chromosome and from the 5' end to the 3' end of transcribed genes. The 3' enrichment suggests that dosage compensation in *Drosophila* may act at the level of transcription elongation [54,55].

The earlier notion that MOF, a global activator of transcription, was harnessed to balance the X-chromosomal monosomy in male *Drosophila* is supported by the fact that the H4K16-specific acetyltransferase activity has been conserved during evolution, although its biological function has not [56,57]. MOF (KAT8) is the best-studied member of the evolutionarily conserved family of MYST acetyltransferases (MOZ (monocytic leukemia zinc finger protein), YBF2/SAS3 (something about silencing 3), SAS2 and TIP60 (60 kDa Tat-interactive protein)). To the best of our knowledge, mammalian MOF is not involved in dosage compensation, but in regulating gene expression in more specific ways and in maintaining genome stability. Knock-down of human MOF impairs the signaling of DNA damage via the ATM pathway in response to double-strand breaks, causing increased cell death and a loss of the cell-cycle checkpoint response [58]. Mouse MOF is essential for oogenesis and embryogenesis [59]. Loss of H4K16ac is a cancer hallmark [60] and MOF is deregulated in a number of diseases [61,62].

As in *Drosophila*, mammalian MOF resides in several distinct complexes. These include the MOF-MLL1-NSL complex, which is required for the expression of the Hox 9a gene [63]; a complex containing the homologs of the *Drosophila* MSL3 and MSL1 that contributes to global



H4K16 acetylation [64,65]; and a complex most closely related to the *Drosophila* NSL complex [52], containing human NSL1 (MSL1v1) and PHF20 (PHD finger protein 20, the homolog of MBD-R2), in addition to other NSL protein homologs. This complex has attracted particular attention as it is not only responsible for the majority of H4K16ac in human cells [66], but also acetylates p53 at lysine 120 (K120) [66,67]. p53 in which K120 is mutated can no longer trigger the apoptotic pathway, yet its role in the cell-cycle checkpoint is not impaired. Evidently, the substrate specificity of human MOF and the physiological processes in which it is involved are largely determined by the molecular context of the acetyltransferase, defined by the composition of the different complexes. In *Drosophila*, however, one of the complexes has been adapted to serve the goal of balancing the genome for dosage compensation.

Negotiation for small effects

Although the mechanisms through which aneuploidies are compensated for are still mysterious, a number of overarching principles have emerged during recent years,

mainly through studies of the X-chromosome monosomies. First, there is no simple switch for 'two-fold up' or 'two-fold down'. Optimal expression levels are negotiated by opposing principles. The X-chromosomal expression in hermaphrodite *C. elegans* results from integration of a global, twofold increase in expression in both sexes and a different counteracting hermaphrodite-specific principle, which halves the expression again (Figure 1c).

The first genome-wide comparison of copy number and transcription in *Drosophila* revealed that a local or chromosomal hemizygosity is compensated for by the integration of at least two different mechanisms: an approximately 1.5-fold compensation can be attributed to general buffering or feedback effects, whereas the remaining compensation is contributed by the evolution of a feed-forward mechanism involving a dedicated dosage-compensation complex [7] (Figure 3b). Furthermore, the twofold activation in male *Drosophila* is a composite of a much larger stimulation, which is opposed by a repressive principle (Figure 3a). We therefore envisage that adjustment of the optimal gene expression levels may be a consequence of negotiation between a number of counteracting activating and repressing principles (Figure 3c). The complex and layered organization of chromatin appears to us as an advanced equalizer with many levers to allow optimal tuning of the transcription melody.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft through SFB-TR5 and the Gottfried-Wilhelm-Leibniz Program. We thank T Straub, C Regnard and T Fauth for comments that improved the manuscript. CF is a fellow of the International Max-Planck Research School in Munich.

Published: 26 August 2010

References

- Veitia RA, Bottani S, Birchler JA: Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. *Trends Genet* 2008, **24**:390-397.
- Schinzl A: *Catalogue of Unbalanced Chromosome Aberrations in Man*. Berlin: Walter de Gruyter; 2001.
- Bannon JH, McGee MM: Understanding the role of aneuploidy in tumorigenesis. *Biochem Soc Trans* 2009, **37**:910-913.
- Birchler JA: Reflections on studies of gene expression in aneuploids. *Biochem J* 2010, **426**:119-123.
- Deng X, Disteche CM: Genomic responses to abnormal gene dosage: the X chromosome improved on a common strategy. *PLoS Biol* 2010, **8**:e1000318.
- Zhang Y, Oliver B: Dosage compensation goes global. *Curr Opin Genet Dev* 2007, **17**:113-120.
- Zhang Y, Malone JH, Powell SK, Periwé V, Spana E, Macalpine DM, Oliver B: Expression in aneuploid *Drosophila* s2 cells. *PLoS Biol* 2010, **8**:e1000320.
- FitzPatrick DR: Transcriptional consequences of autosomal trisomy: primary gene dosage with complex downstream effects. *Trends Genet* 2005, **21**:249-253.
- Ait Yahya-Graison E, Aubert J, Dauphinot L, Rivals I, Prieur M, Golfier G, Rossier J, Personnaz L, Creau N, Bléhaut H, Robin S, Delabar JM, Potier MC: Classification of human chromosome 21 gene-expression variations in Down syndrome: impact on disease phenotypes. *Am J Hum Genet* 2007, **81**:475-491.
- Makarevitch I, Phillips RL, Springer NM: Profiling expression changes caused by a segmental aneuploid in maize. *BMC Genomics* 2008, **9**:7.
- Stenberg P, Lundberg LE, Johansson AM, Ryden P, Svensson MJ, Larsson J: Buffering of segmental and chromosomal aneuploidies in *Drosophila melanogaster*. *PLoS Genet* 2009, **5**:e1000465.
- Ohlsson R: Genetics. Widespread monoallelic expression. *Science* 2007, **318**:1077-1078.
- Lucchesi JC, Kelly WG, Panning B: Chromatin remodeling in dosage compensation. *Annu Rev Genet* 2005, **39**:615-651.
- Gupta V, Parisi M, Sturgill D, Nuttall R, Doctolero M, Dudko OK, Malley JD, Eastman PS, Oliver B: Global analysis of X-chromosome dosage compensation. *J Biol* 2006, **5**:3.
- Nguyen DK, Disteche CM: Dosage compensation of the active X chromosome in mammals. *Nat Genet* 2006, **38**:47-53.
- Lin H, Gupta V, Vermilyea MD, Falciani F, Lee JT, O'Neill LP, Turner BM: Dosage compensation in the mouse balances up-regulation and silencing of X-linked genes. *PLoS Biol* 2007, **5**:e326.
- Graves JA, Disteche CM: Does gene dosage really matter? *J Biol* 2007, **6**:1.
- Arnold AP, Itoh Y, Melamed E: A bird's-eye view of sex chromosome dosage compensation. *Annu Rev Genomics Hum Genet* 2008, **9**:109-127.
- Mank JE: The W, X, Y and Z of sex-chromosome dosage compensation. *Trends Genet* 2009, **25**:226-233.
- Charlesworth B: The evolution of chromosomal sex determination and dosage compensation. *Curr Biol* 1996, **6**:149-162.
- Marin I, Siegal ML, Baker BS: The evolution of dosage-compensation mechanisms. *Bioessays* 2000, **22**:1106-1114.
- Charlesworth D, Charlesworth B, Marais G: Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 2005, **95**:118-128.
- Larsson J, Meller VH: Dosage compensation, the origin and the afterlife of sex chromosomes. *Chromosome Res* 2006, **14**:417-431.
- Marshall Graves JA, Peichel CL: Are homologies in vertebrate sex determination due to shared ancestry or to limited options? *Genome Biol* 2010, **11**:205.
- Birchler JA, Bhadra U, Bhadra MP, Auger DL: Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. *Dev Biol* 2001, **234**:275-288.
- Charlesworth B: Model for evolution of Y chromosomes and dosage compensation. *Proc Natl Acad Sci USA* 1978, **75**:5618-5622.
- Mukherjee AS, Beermann W: Synthesis of ribonucleic acid by the X-chromosomes of *Drosophila melanogaster* and the problem of dosage compensation. *Nature* 1965, **207**:785-786.
- Straub T, Gilfillan GD, Maier VK, Becker PB: The *Drosophila* MSL complex activates the transcription of target genes. *Genes Dev* 2005, **19**:2284-2288.
- Hamada FN, Park PJ, Gordadze PR, Kuroda MI: Global regulation of X chromosomal genes by the MSL complex in *Drosophila melanogaster*. *Genes Dev* 2005, **19**:2289-2294.
- Gelbart ME, Kuroda MI: *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* 2009, **136**:1399-1410.
- Grimaud C, Becker PB: Form and function of dosage-compensated chromosomes - a chicken-and-egg relationship. *BioEssays* 2010, **32**:709-717.
- Straub T, Becker PB: Dosage compensation: the beginning and end of generalization. *Nat Rev Genet* 2007, **8**:47-57.
- Straub T, Becker PB: DNA sequence and the organization of chromosomal domains. *Curr Opin Genet Dev* 2008, **18**:175-180.
- Robinson PJ, Fairall L, Huynh VA, Rhodes D: EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci USA* 2006, **103**:6506-6511.
- Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL: Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 2006, **311**:844-847.
- Oliver B: Sex, dose, and equality. *PLoS Biol* 2007, **5**:e340.
- Payer B, Lee JT: X chromosome dosage compensation: how mammals keep the balance. *Annu Rev Genet* 2008, **42**:733-772.
- Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT: Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 2008, **322**:750-756.
- Maenner S, Blaud M, Fouillen L, Savoye A, Marchand V, Dubois A, Sanglier-Cianférani S, Van Dorsselaer A, Clerc P, Avner P, Visvikis A, Branlant C: 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS Biol*, **8**:e1000276.
- Jans J, Gladden JM, Ralston EJ, Pickle CS, Michel AH, Pferdehirt RR, Eisen MB, Meyer BJ: A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. *Genes Dev* 2009, **23**:602-618.

41. Cleard F, Spierer P: **Position-effect variegation in *Drosophila*: the modifier Su(var)3-7 is a modular DNA-binding protein.** *EMBO Rep* 2001, **2**:1095-1100.
42. Spierer A, Seum C, Delattre M, Spierer P: **Loss of the modifiers of variegation Su(var)3-7 or HP1 impacts male X polytene chromosome morphology and dosage compensation.** *J Cell Sci* 2005, **118**:5047-5057.
43. Spierer A, Begeot F, Spierer P, Delattre M: **SU(VAR)3-7 links heterochromatin and dosage compensation in *Drosophila*.** *PLoS Genet* 2008, **4**:e1000066.
44. Badenhorst P, Voas M, Rebey I, Wu C: **Biological functions of the ISWI chromatin remodeling complex NURF.** *Genes Dev* 2002, **16**:3186-3198.
45. Deuring R, Fanti L, Armstrong JA, Sarte M, Papoulas O, Prestel M, Daubresse G, Verardo M, Moseley SL, Berloco M, Tsukiyama T, Wu C, Pimpinelli S, Tamkun JW: **The ISWI chromatin remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo.** *Mol Cell* 2000, **5**:355-365.
46. Corona DF, Siriaco G, Armstrong JA, Snarskaya N, McClymont SA, Scott MP, Tamkun JW: **ISWI regulates higher-order chromatin structure and histone H1 assembly in vivo.** *PLoS Biol* 2007, **5**:e232.
47. Buscaino A, Kocher T, Kind JH, Holz H, Taipale M, Wagner K, Wilm M, Akhtar A: **MOF-regulated acetylation of MSL-3 in the *Drosophila* dosage compensation complex.** *Mol Cell* 2003, **11**:1265-1277.
48. Morales V, Straub T, Neumann MF, Mengus G, Akhtar A, Becker PB: **Functional integration of the histone acetyltransferase MOF into the dosage compensation complex.** *EMBO J* 2004, **23**:2258-2268.
49. Robinson PJ, Rhodes D: **Structure of the '30 nm' chromatin fibre: a key role for the linker histone.** *Curr Opin Struct Biol* 2006, **16**:336-343.
50. Akhtar A, Becker PB: **Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*.** *Mol Cell* 2000, **5**:367-375.
51. Prestel M, Feller C, Straub T, Mitlöhner H, Becker PB: **The activation potential of MOF is constrained for dosage compensation.** *Mol Cell* 2010, **38**:815-826.
52. Mendjan S, Taipale M, Kind J, Holz H, Gebhardt P, Schelder M, Vermeulen M, Buscaino A, Duncan K, Mueller J, Wilm M, Stunnenberg HG, Saumweber H, Akhtar A: **Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*.** *Mol Cell* 2006, **21**:811-823.
53. Bienz M: **The PHD finger, a nuclear protein-interaction domain.** *Trends Biochem Sci* 2006, **31**:35-40.
54. Gilfillan GD, Straub T, de Wit E, Greil F, Lamm R, van Steensel B, Becker PB: **Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex.** *Genes Dev* 2006, **20**:858-870.
55. Alekseyenko AA, Peng S, Larschan E, Gorchakov AA, Lee OK, Kharchenko P, McGrath SD, Wang CI, Mardis ER, Park PJ, Kuroda MI: **A sequence motif within chromatin entry sites directs MSL establishment on the *Drosophila* X chromosome.** *Cell* 2008, **134**:599-609.
56. Rea S, Xouri G, Akhtar A: **Males absent on the first (MOF): from flies to humans.** *Oncogene* 2007, **26**:5385-5394.
57. Li X, Dou Y: **New perspectives for the regulation of acetyltransferase MOF.** *Epigenetics* 2010, **5** DOI:10.4161/epi.5.3.11372
58. Gupta A, Sharma GG, Young CS, Agarwal M, Smith ER, Paull TT, Lucchesi JC, Khanna KK, Ludwig T, Pandita TK: **Involvement of human MOF in ATM function.** *Mol Cell Biol* 2005, **25**:5292-5305.
59. Gupta A, Guerin-Peyrou TG, Sharma GG, Park C, Agarwal M, Ganju RK, Pandita S, Choi K, Sukumar S, Pandita RK, Ludwig T, Pandita TK: **The mammalian ortholog of *Drosophila* MOF that acetylates histone H4 lysine 16 is essential for embryogenesis and oncogenesis.** *Mol Cell Biol* 2008, **28**:397-409.
60. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Pérez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, Esteller M: **Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer.** *Nat Genet* 2005, **37**:391-400.
61. Pfister S, Rea S, Taipale M, Mendrzyk F, Straub B, Itrich C, Thuerigen O, Sinn HP, Akhtar A, Lichter P: **The histone acetyltransferase hMOF is frequently downregulated in primary breast carcinoma and medulloblastoma and constitutes a biomarker for clinical outcome in medulloblastoma.** *Int J Cancer* 2008, **122**:1207-1213.
62. Kapoor-Vazirani P, Kagey JD, Powell DR, Vertino PM: **Role of hMOF-dependent histone H4 lysine 16 acetylation in the maintenance of TMS1/ASC gene activity.** *Cancer Res* 2008, **68**:6810-6821.
63. Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, Wysocka J, Allis CD, Chait BT, Hess JL, Roeder RG: **Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF.** *Cell* 2005, **121**:873-885.
64. Smith ER, Cayrou C, Huang R, Lane WS, Cote J, Lucchesi JC: **A human protein complex homologous to the *Drosophila* MSL complex is responsible for the majority of histone H4 acetylation at lysine 16.** *Mol Cell Biol* 2005, **25**:9175-9188.
65. Taipale M, Rea S, Richter K, Vilar A, Lichter P, Imhof A, Akhtar A: **hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells.** *Mol Cell Biol* 2005, **25**:6798-6810.
66. Li X, Wu L, Corsa CA, Kunkel S, Dou Y: **Two mammalian MOF complexes regulate transcription activation by distinct mechanisms.** *Mol Cell* 2009, **36**:290-301.
67. Sykes SM, Mellert HS, Holbert MA, Li K, Marmorstein R, Lane WS, McMahon SB: **Acetylation of the p53 DNA-binding domain regulates apoptosis induction.** *Mol Cell* 2006, **24**:841-851.
68. Fauth T, Muller-Planitz F, König C, Straub T, Becker PB: **The DNA binding CXC domain of MSL2 is required for faithful targeting the Dosage Compensation Complex to the X chromosome.** *Nucleic Acids Res* 2010, **38**:3209-3221.
69. Morra R, Smith ER, Yokoyama R, Lucchesi JC: **The MLE subunit of the *Drosophila* MSL complex uses its ATPase activity for dosage compensation and its helicase activity for targeting.** *Mol Cell Biol* 2008, **28**:958-966.
70. Oh H, Parrott AM, Park Y, Lee CG: **Regulation of inter- and intramolecular interaction of RNA, DNA, and proteins by MLE.** *Methods Mol Biol* 2010, **587**:303-326.

doi:10.1186/gb-2010-11-8-216

Cite this article as: Prestel M, et al.: Dosage compensation and the global re-balancing of aneuploid genomes. *Genome Biology* 2010, **11**:216.

3.3 The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset.

Christian Feller, Matthias Prestel, Holger Hartmann, Tobias Straub,
Johannes Söding and Peter B. Becker

3.3.1 Summary, significance and own contribution

Summary, significance and discussion with recent literature

The NSL complex is found at a subset of active genes but what determines whether a gene is bound and activated by the complex was unknown. In this study, we generated new and analysed existing genomic data sets to unveil the determinants that target the complex and define whether it engages in transcriptional activation. We found that the NSL complex primarily targets the promoters of active housekeeping genes. There, it co-localises with WDS, the chromatin remodeler subunit NURF301, the H3.K4 methyltransferase trithorax and the interband protein chromator. Importantly, only a subset of the genes that associate with the NSL complex are regulated by it. Comparing ChIP-chip data of chromatin regulators and promoter DNA sequences between regulated and non-regulated genes revealed that the set of NSL-activated promoters are enriched for the promoter DNA motif ‘Ohler 5’ and depleted for the insulators CP190 and BEAF as well as the heterochromatin protein 1c (HP1c). Together, these results show that the capacity of the NSL complex to activate transcription is highly context dependent. Moreover, our study suggests that not only tissue-specific genes employ dedicated transcription factors, but that housekeeping genes are also regulated by distinct sets of co-activators.

Our observation that NURF301 and trithorax co-localise with the NSL complex was surprising at the time because these proteins were traditionally studied for their roles in regulating very restricted gene sets during specific developmental programs (Ringrose and Paro, 2004; Badenhorst et al., 2005). However, there is accumulating evidence that, under specific conditions, trithorax and its mammalian homologue MLL1 interact with the NSL complex (Petruk et al., 2001; Dou et al., 2005; Zhao et al., 2013; Tie et al., 2014). In addition, the bromodomain of the mammalian NURF301 homolog, BPTF, has been reported to bind H4.K16ac (Ruthenburg et al., 2011).

Recent studies confirmed our finding of NSL complex binding to housekeeping genes in *Drosophila* cells (Lam et al., 2012) and reported the conservation of this binding mode also in mouse cells (Chelmicki et al., 2014; Ravens et al., 2014). Interestingly, a deeper comparison of these articles also reveals some noteworthy differences. First, while we reported that only a minor subset of NSL bound genes are down-regulated after ablating either NSL1 or MBD-R2, Akhtar and colleagues described that a major fraction of NSL occupied genes are regulated by the complex, as assessed by measuring polymerase II occupancy on target and non-target gene promoters upon depleting NSL1 or NSL3 (Lam et al., 2012). Furthermore, Akhtar and co-workers reported that the core promoter motif ‘DRE’ correlates best with the capacity of the NSL complex to regulate transcription, but our analysis suggests that while the binding ‘strength’ of NSL1 correlates best with this motif (‘DRE’), transcriptional regulation by the complex is rather associated with motif ‘Ohler 5’. These discrepancies can be, at least in part, attributed to the different approaches to evaluate the effect on transcription upon NSL ablation. While we monitored directly the steady-state mRNA levels, Akhtar and colleagues used diminished ChIP occupancy of the RNA polymerase II subunit Rpb3 as a proxy

for transcriptional regulation. It is conceivable that reduced polymerase II loading at promoters may not always translate into a decrease of mRNA production. Likewise, monitoring steady-state mRNA levels may be too insensitive in cases where newly generated transcripts are outnumbered by a large pool of pre-existing mRNA molecules. An alternative and complementary approach would be to analyse the polymerase II isoform that is phosphorylated on serine 5 at its CTD along gene bodies, which has been documented to show an overall good correlation with gene activity measured by bulk mRNA levels (Kharchenko et al., 2011; Regnard et al., 2011).

Second, while the *Drosophila* NSL complex primarily binds promoters (Feller et al., 2012; Lam et al., 2012), a substantial fraction of mouse NSL binding events was additionally detected at enhancers (Chelmicki et al., 2014). Interestingly, while Akhtar and colleagues reported that the mouse NSL complex regulates key pluripotency factors in addition to housekeeping genes and that most transcription regulatory potential is observed from NSL sites at enhancers (Chelmicki et al., 2014), Tora and colleagues only observed minor intergenic binding of mouse NSL1 and no deregulation of pluripotency genes (Ravens et al., 2014). This discrepancy may be in part attributed to monitoring and ablating different components of the NSL complex (NSL1 in Ravens et al. vs. NSL3 and MCRS2 in Chelmicki et al.).

An important recent structure-function study provided strong evidence that the interaction between WDR5 (human homolog to the fly WDS), which is a component of many multi-protein complexes including the MLL1 complex and the ATAC complex (Suganuma et al., 2008; Shilatifard, 2012), is also a central constituent of the NSL complex that tethers NSL1 to NSL2 and is essential for recruitment of the complex to its target genes (Dias et al., 2014). This study also shows that WDS uses the same regions for interacting with NSL1 and NSL2 as it does for its interaction within the MLL1 complex. This finding provides strong structural support that WDR5 is a shared component of the NSL and MLL complex rather than tethering the acetyltransferase and methyltransferase complexes to a super-complex, as it has been suggested before (Dou et al., 2005; Li et al., 2009).

Own contribution

Prof. Peter Becker and I conceived the project and wrote the manuscript. I performed all genomic and cell biological experiments, analysed the data and prepared all figures. I initiated the collaboration with Dr. Johannes Söding and Holger Hartmann, who performed the core promoter analysis. Dr. Matthias Prestel developed and characterised the NSL1 antibody and conducted most of the reporter gene experiments. Dr. Tobias Straub supervised the bioinformatic analysis and provided important scripts.

3.3.2 Published manuscript

The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset

Christian Feller¹, Matthias Prestel¹, Holger Hartmann², Tobias Straub¹,
Johannes Söding² and Peter B. Becker^{1,*}

¹Adolf-Butenandt-Institute and Center for Integrated Protein Science of the Ludwig-Maximilians-University, Schillerstraße 44, 80336 München, Germany and ²Gene Center Munich, Department of Chemistry and Biochemistry, Ludwig-Maximilians-University, Feodor-Lynen-Straße 25, 81377 Munich, Germany

Received September 5, 2011; Revised September 27, 2011; Accepted September 28, 2011

ABSTRACT

The MOF (males absent on the first)-containing NSL (non-specific lethal) complex binds to a subset of active promoters in *Drosophila melanogaster* and is thought to contribute to proper gene expression. The determinants that target NSL to specific promoters and the circumstances in which the complex engages in regulating transcription are currently unknown. Here, we show that the NSL complex primarily targets active promoters and in particular housekeeping genes, at which it colocalizes with the chromatin remodeler NURF (nucleosome remodeling factor) and the histone methyltransferase Trithorax. However, only a subset of housekeeping genes associated with NSL are actually activated by it. Our analyses reveal that these NSL-activated promoters are depleted of certain insulator binding proteins and are enriched for the core promoter motif 'Ohler 5'. Based on these results, it is possible to predict whether the NSL complex is likely to regulate a particular promoter. We conclude that the regulatory capacity of the NSL complex is highly context-dependent. Activation by the NSL complex requires a particular promoter architecture defined by combinations of chromatin regulators and core promoter motifs.

INTRODUCTION

Eukaryotic organisms consist of a diversified set of highly specialized cells. Their individual identities are determined by the appropriate expression of cell-specific genes while a battery of genes that are expressed in all cells maintain general ('housekeeping') functions. Gene expression at

the transcriptional level is governed by an intricate interplay between transcription regulators and local chromatin organization. In general, the packaging of genomes into chromatin brings about a default state of repression, as nucleosome assembly constantly competes with transcription factors for promoter binding sites. Overcoming this repression requires a concerted action of various chromatin-modifying principles. These include ATP-dependent nucleosome remodeling factors, which are targeted to specific loci by DNA-bound proteins and post-translational histone marks where they reorganize nucleosomes to facilitate transcription (1). An example for such an activity in *Drosophila melanogaster* is NURF (nucleosome remodeling factor), whose large regulatory subunit, NURF301, interacts with a diversity of transcription factors and methyl marks on lysine 4 of histone H3 (H3K4me3) (2,3) (and references therein). NURF has also been reported to bind to acetylated lysine 16 of histone H4 (H4K16ac) (2), a nucleosome modification that prevents nucleosome–nucleosome interactions that promote the folding of the nucleosomal fiber into more compact structures. The acetyltransferase MOF (males absent on the first) is a major enzyme responsible for this modification in both, *Drosophila* and mammalian cells (4,5).

MOF is best known for its key role in the *Drosophila* dosage compensation process. It is a subunit of the dosage compensation complex [DCC, also known as male-specific lethal (MSL) complex], which brings about the 2-fold transcriptional activation of genes on the single male X chromosome to equalize expression with the corresponding genes transcribed from the two female X chromosomes (6). The DCC is constituted only in male flies and the five protein components, MSL1, MSL2, MSL3, maleless (MLE) and MOF, as well as the non-coding *roX* RNAs are essential for male viability. According to the current model, the DCC recruits MOF to the transcribed regions of X-chromosomal genes. Subsequent acetylation of

*To whom correspondence should be addressed. Tel: 089 2180 75 427; Fax: 089 2180 75 425; Email: pbecker@med.uni-muenchen.de

H4K16 renders chromatin more accessible and potentially facilitates transcriptional elongation (7,8).

With the exception of MSL2, all DCC protein subunits are also expressed in female flies, and therefore also serve more general, yet barely understood functions (9). For example, the acetyltransferase MOF appears to be involved in more global transcription regulation as it has recently been found in an alternative complex together with MCRS2, the WD40-repeat protein WDS (will-die-slowly), NSL1, NSL2, NSL3 and the plant homeo domain (PHD) protein MBD-R2 (10–12). With reference to the dosage compensation ‘MSL complex’, this alternative MOF-containing assembly was termed ‘NSL complex’ (for ‘non-specific lethal’), as its subunits are essential in both sexes (10). The incorporation of MOF into either the DCC or the NSL complex is determined by association of MOF with the PEHE domains of the respective MSL1 or NSL1 subunits (10). Genome-wide mapping by chromatin immunoprecipitation (ChIP) coupled to DNA microarrays (ChIP-chip) identified MOF binding sites at many, but not all active promoters in male and female cells (13). Subsequent studies revealed that MBD-R2 colocalizes with MOF at many active promoters in both sexes, suggesting that the NSL complex recruits MOF to these sites (12). This is compatible with a recent ChIP-Seq study (ChIP DNA analyzed by massive parallel sequencing), which found MCRS2 and NSL1 peaks at promoters in mixed-sex 3rd instar larval salivary glands (11).

In male cells the association of MOF with NSL subunits is in competition with its incorporation into the DCC, which redirects it to the transcribed regions of X chromosomal genes (12). However, key aspects of MOF’s targeting in the context of the NSL complex are unclear. What determines the binding of the NSL complex to only a subset of the active promoters? The available data also are ambiguous when it comes to the role of the NSL complex; does it activate or repress target genes, or perhaps both? Ablating the NSL subunit MBD-R2 in male embryonic cells resulted in a reduced expression of many MBD-R2 target genes (12). In contrast, a similar fraction of genes was found up- and downregulated when MBD-R2 and NSL3 were depleted in 3rd instar salivary glands (11).

In this study, we created novel data sets and analyzed existing ones to compare functional interactions of NSL subunits in different developmental tissues to better define the targets of the NSL complex. We systematically explored the common properties of the NSL target

genes, searching for colocalizing chromatin factors and prevalent sequence motifs in target promoters. We traced the NSL complex through monitoring the NSL1 subunit and found that it preferentially binds to promoters of housekeeping genes, which are also approached by the chromatin remodeler NURF and the methyltransferase Trithorax. There, NSL1 binding correlates best with the core promoter element DNA replication-related element (DRE). However, only a defined fraction of NSL1-bound genes are actually regulated by the complex. Those promoters are depleted for insulator proteins and are enriched for the E-box-derived promoter motif ‘Ohler 5’. Our analysis provides a functional classification of housekeeping genes according to their NSL coregulator requirements.

MATERIALS AND METHODS

Generation of the NSL1 antibody

A cDNA fragment corresponding to NSL1 amino acids 1271–1550 was Polymerase Chain Reaction (PCR) amplified from cDNA clone #LP09056 (*Drosophila* Genomics Resource Center; see Table 1) and cloned into the pGEX2TKN. The N-terminally glutathion-S-transferase (GST)-tagged NSL1 fragment was expressed in *Escherichia coli* BL21, purified on glutathione beads and used to raise antibodies in rabbit by a commercial supplier.

RNA interference in S2 cells, immunoblotting and indirect immunofluorescence

Male *Drosophila* S2 cell cultivation and RNA interference (RNAi) were carried out as described before (12). Briefly, 1.5×10^6 cells were incubated with 10 µg dsRNA targeted against NSL1 or GST as a control. Primer sequences used for dsRNA production are listed in Table 1. Cells were harvested after 6 or 7 days and processed for RNA (see below) and protein. For every 10^6 cells, cells were lysed for 10 min in 100 µl of N-buffer [15 mM (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.5, 60 mM KCl, 15 mM NaCl, 0.5 mM ethylene glycol tetraacetic acid pH 8, 0.25% Triton-X, 10 mM sodium butyrate, 1 mM phenylmethanesulfonylfluoride, 0.1 mM Dithiothreitol protease inhibitor cocktail (Roche)] on ice and the chromatin fraction was pelleted by centrifugation. RNA for Affymetrix expression profiling was prepared as described (12). RNA labeling and cDNA hybridization to a *Drosophila*

Table 1. Primer table

Construct	Forward primer sequence	Reverse primer sequence
NSL1 RNAi amplicon 1	TTAATACGACTCACTATAGGGA GCGTC CGAGCTCAAC CTTC	TTAATACGACTCACTATAGGGA CACATGGGTGTGTTTCATTAGTC
NSL1 RNAi amplicon 2	TTAATACGACTCACTATAGGGA GATGTCGCATCAAAGTCAGAGG	TTAATACGACTCACTATAGGGA GACTCGAGAAGAGCTCGCTGAT
GST RNAi amplicon	TTAATACGACTCACTATAGGGAG AATGTCCCCTATACTAG GTTA	TTAATACGACTCACTATAGGGAGA ACGCAT CCAGGCACATTG
NSL1 antibody cloning	CGCTCCATGGCTTTTCATT AAGTCCCCTGGAGCACC	ATTCTAGATTAGATGC GTCTGCTGCGAACACCCTC

Genome GeneChip 2.0 was performed at the Gene Center Affymetrix Microarray Platform (Munich, Germany). Immunoblot analysis and immunofluorescence microscopy (IFM) analysis was performed as described previously (14). The lamin antibody was obtained from H. Saumweber (Berlin) and the MSL1 antibody was described previously (15).

Reporter gene ChIP assay and luciferase reporter assay

The reporter gene ChIP assay and luciferase reporter assay have been described before (12).

Chromatin extraction and immunoprecipitation

Chromatin extraction and immunoprecipitation were previously described (12). Briefly, chromatin extracts from sex-sorted adult flies were prepared and the DNA concentration of the extract was determined. DNA (7.5–15 µg) were used for a single ChIP experiment. Five microliters of anti-NSL1 serum was used in a single IP reaction. After the precipitation and extensive washing, DNA was extracted with phenol/chloroform, ethanol precipitated and further cleaned using the GenElute PCR clean-up kit (SIGMA). DNA was amplified using the whole-genome amplification kit (WGA, SIGMA). Labeling, hybridization to customized high-resolution NimbleGen tiling arrays (comprising the euchromatic part of the entire X chromosome, 5 Mb of 2 L, 2 R and 3 L, respectively, as well as 10 Mb of 3 R) (12), scanning and feature extraction was performed by imaGenes (Berlin).

ChIP-chip data processing

ChIP-chip data analysis was performed using R/Bioconductor (www.r-project.org; www.bioconductor.org). Raw signals of the NimbleGen NSL1 ChIP-chip were normalized and log2-transformed using the ‘vsn’ package (16). IP/input ratios of the modENCODE data were scaled to a mean of zero and a standard deviation of one. Promoter enrichments were calculated by summarizing the probe level signals in a window of 600 bp centered at the transcriptional start site (TSS) (FlyBase release 5.22). Promoter binding was classified based on the bimodal distribution of binding values, where genes within the population of lower values were considered ‘unbound’ and genes within the population of higher values were considered ‘bound’. Alternatively, ‘bound’ were selected based on the *fdr* values from the ‘locfdr’ package applied on the promoter binding values with a *fdr* cutoff of <0.2. The results are robust to several normalization methods and promoter window definitions.

Genes were classified ‘active’ when (i) their Affymetrix expression value exceeded four (see below) and (ii) RNA polymerase II [modENCODE profile (17)] was classified as ‘bound’ on their promoters. A similar result was obtained using genes which are ‘bound’ (modeled on the bimodal distribution of the averaged binding along the transcribed region) by the elongating polymerase [serine 2 phosphorylated RNA polymerase II, data from (18)].

Promoters were classified as ‘peaked’, ‘broad’ and ‘weak peak’ promoters according to Hoskins *et al.* (19) and Ni *et al.* (20). Hierarchical cluster analysis of the promoter

binding pattern was carried out using the R package ‘hclust’ and the ‘complete’ or ‘ward’ clustering approach as indicated in the figure legends.

All available modENCODE chromatin ChIP-chip data sets were screened for factors, which are enriched at promoter locations (by March 2011). After initial data quality assessment probe level binding was assessed for promoter probes (broad: ±300 bp centered at TSS; narrow: ±100 bp centered at TSS; upstream-biased: -300 - TSS- +100 bp), transcriptional termination (TT) sites (broad: ±300 bp centered at TT; narrow: ±100 bp centered at TT; downstream-biased: -100 - TT+ 300 bp), gene probes (probes corresponding to annotated genes without promoter and termination probes) and intergenic probes (defined as probes not found in previous groups). Only ChIP-chip data sets with a clear enrichment for promoter probes relative to gene, intergenic and termination probes were selected for this study.

Transcriptome data analysis

Transcriptome data analysis was conducted as described previously (12). Briefly, raw signals were normalized, summarized and log2-transformed using the ‘gcrma’ package. Significant change of gene expression was calculated applying the ‘locfdr’ package on a ‘sam’ statistics using a cutoff of *fdr* <0.35. Alternatively, genes with $\log_2(\text{NSL1 RNAi-GST RNAi}) < (-1)$ were considered ‘down-regulated’. The results are robust to various parameters in data analysis, as assessed by choosing varying thresholds. All expression data set values are log2-normalized with a theoretical dynamic range of 2^{exp16} (Affymetrix.com).

Housekeeping gene definition

Affymetrix expression data sets of 40 different *Drosophila* tissues [GSE7763, (21)] were processed as described above for the NSL1 transcriptome data set. For every gene, the standard deviation was calculated across all 40 samples (gene variation index). Filtering for active genes, the distribution of standard deviations resulted in two major populations with the best discrimination at a standard deviation of ~1.5 (Supplementary Figure S9A and B). Consequently, genes with a gene variation index <1.5 were considered housekeeping genes and genes with a gene variation index >1.5 were considered differentially regulated genes. The results are robust to different applied thresholds. In an alternative analysis (presented in Supplementary Figure S2E), we took the more stringent call for housekeeping gene function according to the classification of Weber and Hurst (22). Here, active genes which belong to either the ‘tau’ class or to the ‘breadth’ class were considered housekeeping genes.

ChIP-Seq data analysis

NSL1 ChIP-Seq and corresponding input data sets (11) were obtained from the ArrayExpress repository (E-MTAB-214). Sequence reads were mapped to the *Drosophila melanogaster* genome (dm3) using bowtie (23). Uninformative reads and read anomalies were filtered out using the R package ‘SPP’ (24), resulting in 7840131

unique NSL1 ChIP reads and 6094163 unique input reads. Peaks were identified using SPP with the following parameters: ‘tag-wtd’ method, $\text{fdr} = 0.01$, minimal distance between detected peaks = 100 bp. The input data was used to determine statistical significance of NSL1 peaks, resulting in the ‘peak score’.

Core promoter motif analysis

We used the 10 promoter motifs described by Ohler and colleagues (25) to analyze promoter motif occurrences. For every motif a log-odds weight matrix description P of the binding sites is given, which was used to calculate a motif score for a specific sequence. It ranges between zero and one and measures how similar a binding site is to the consensus. In a first step, the log-odds score for the consensus site L_C is determined by

$$L_C = \sum_{i=1}^w \max\{P_{ib} : b \in \{A,C,G,T\}\},$$

where w is the motif length. The motif score given a specific binding site starting at position k in sequence X is calculated by

$$\text{motif score}(X,k) = \frac{1}{L_C} \sum_{i=1}^w P_{iX_{k+i-1}}$$

The motif score is the ratio of the log-odds score of the site at position k to the log-odds score of the consensus site. The motif score for the entire sequence X is given by the highest motif score in the sequence:

$$\text{motif score}(X) = \max_k \{\text{motif score}(X,k)\}$$

For the analyses, we used a threshold of motif score >0.3 to consider a binding site as functional. The *de novo* sequence analysis algorithm will be reported elsewhere (Hartmann and Soeding, manuscript in preparation).

RESULTS

NSL1 colocalizes with MBD-R2 at many active promoters

The genomic interaction profile of MOF differs in adult male and female flies, reflecting its incorporation into the male-specific DCC and the general NSL complex (11,12). We previously monitored the MBD-R2 distribution in adult male and female flies but could not detect any significant difference (12). Since MBD-R2 is the only NSL complex protein which may interact with DCC members (10) we sought to compare the genome-wide binding pattern of the NSL complex with the potential core subunit of the complex, NSL1. In order to compare the NSL1 interactions in the genomes of adult male and female flies, an antibody was raised against NSL1 and

its specificity confirmed combining RNAi with subsequent detection by indirect immunofluorescence microscopy (IFM) and immunoblotting (Supplementary Figure S1 and see below). The antibody was then used for ChIP-chip experiments, where NSL1 was precipitated from chromatin preparations from hand-sorted adult male and female flies and the associated DNA was amplified and hybridized to high-resolution DNA tiling microarrays representing the X chromosome and an equivalent amount of the autosomes. The binding profile in male and female flies did not show any significant differences (Supplementary Figure S2A). In addition, NSL1 was found at the same loci as MBD-R2 (Supplementary Figure S2B), in agreement with the results of the biochemical definition of both proteins as ‘NSL’ complex subunits (10–12). The ChIP-chip profiling suggested that NSL1 globally binds target loci independent of the fly sex, confirming previous ChIP-qPCR analyses at selected loci (11).

Re-examination of the previously published NSL1 ChIP-Seq profiles, which had been generated from salivary glands of mixed-sex third instar larvae (11), revealed a systematic enrichment of NSL1 peaks at RNA polymerase II—promoters relative to genes transcribed by RNA polymerases I and III (Table 2). Applying a superior peak calling algorithm (24) to these data identified the majority of NSL1 binding events within a window of 200 base pairs (bp) around the annotated TSS (Figure 1A), implicating the NSL complex in transcriptional initiation.

In order to avoid the heterogeneous salivary gland tissue, which impedes a comparison of NSL binding with the transcriptional activity and with other known promoter binding factors, an NSL1 ChIP-chip profile was generated from *Drosophila* S2 cells. These cells are commonly used in the chromatin community because they provide a homogeneous biological material, a fact that allows comparing our data to other published genomic data sets, such as the comprehensive collection of chromatin factors and histone modifications generated by the modENCODE consortium with a similar ChIP-chip strategy (17).

The newly generated NSL1 ChIP-chip profile correlated well with our previously published MBD-R2 profile (12) as well as with the MBD-R2 profile generated by the modENCODE consortium using a different antibody (Supplementary Figure S2C). Therefore, in the following we subsume the individual NSL1 and MBD-R2 profiles as the ‘NSL complex’ binding, unless stated otherwise. We related the NSL complex binding at promoters with the transcriptional activity of the corresponding genes, using the ChIP-chip profile of the elongating polymerase as a direct readout for active transcription (18). The NSL complex binds active genes with high preference, but only a subset of ~60–70% (depending on the threshold) (Figure 1B, left). A similar result was obtained when

Table 2. NSL1 ChIP-Seq peaks mapped to transcript type

Transcript type	MiRNA	mRNA	ncRNA	rRNA	snoRNA	snRNA	tRNA
Number annotated transcripts	194	22 765	189	160	249	47	314
Number NSL1 peaks mapped to transcript TSS	0	4302	14	0	1	1	5
Fraction NSL1-bound transcripts rel. to all transcripts	0	18.9	7.4	0	0.4	2.1	1.6

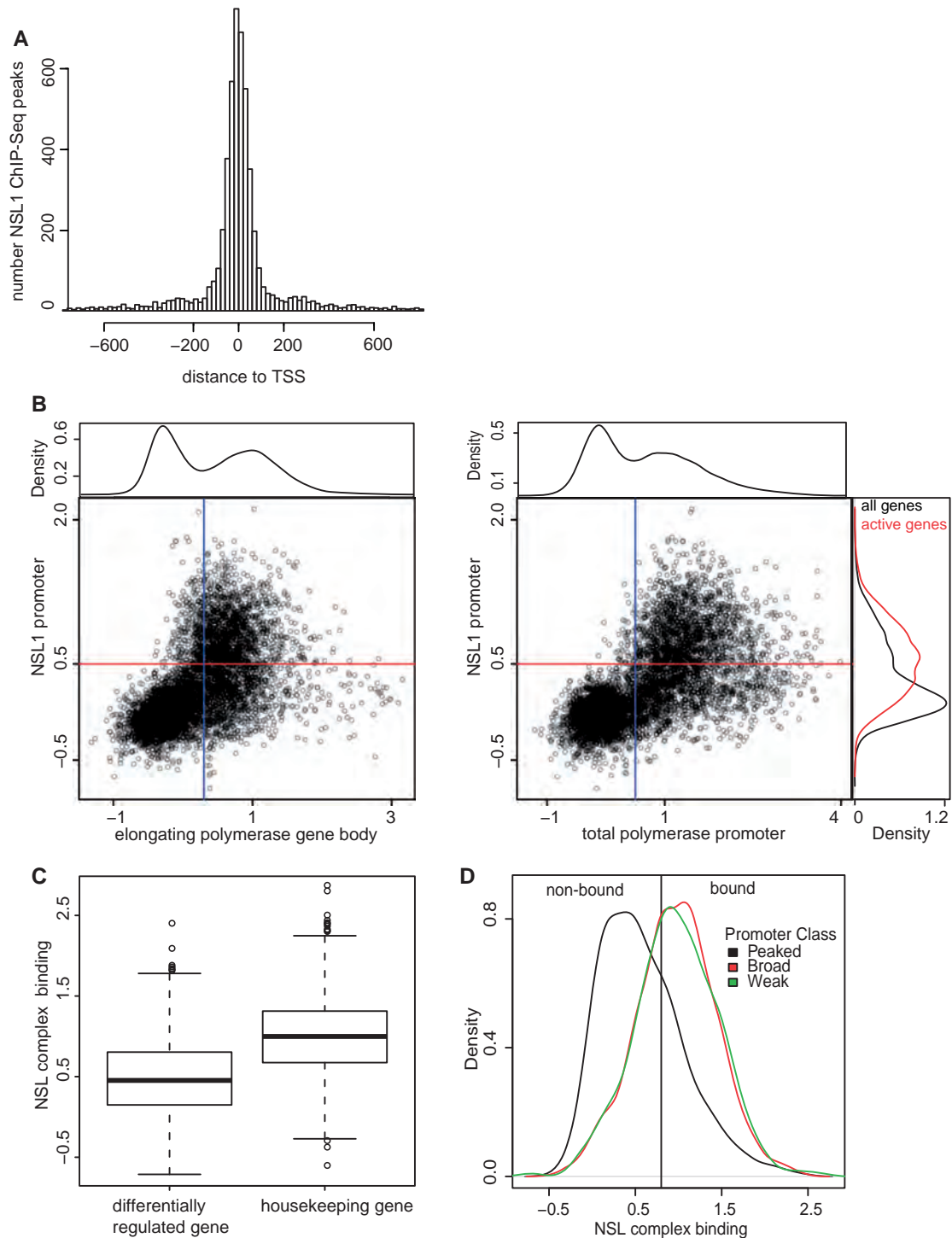


Figure 1. The MOF-containing NSL complex is enriched at promoters of most housekeeping genes. (A) NSL1 peaks map close to the TSS. The histogram displays the distance between the summit of NSL1 ChIP-Seq (11) peak and the closest TSS. Refinement of NSL1 positions relative to the TSS compared to (11) was achieved by using an improved peak calling algorithm [SPP (24) and more precisely mapped TSSs (56)]. (B) NSL1 prevalently binds active gene promoters. Scatter plots of NSL1 promoter binding versus (i) elongating polymerase at all genes (left) (18), (ii) total polymerase promoter occupancy at all genes (right) [modENCODE (17)]. Significant binding cutoffs of NSL1 (red) and polymerase (blue) are indicated. The density plots on the top of each histogram depict the signal distribution of the elongating polymerase (left) and the total polymerase (right). The density plot to the right indicates the NSL1 promoter signal distribution for all genes (black) and active genes [red, based on the elongating polymerase II gene signal (18)], respectively. (C) The NSL complex preferentially associates with housekeeping genes (Welch two sample t -test, $P < 2.2 \times 10^{-16}$). The boxplot depicts NSL1 binding at differentially regulated and housekeeping genes (for categorization, see 'Materials and Methods' section). A similar result is gained using the available MBD-R2 ChIP-chip data sets (12). An alternative, more stringent categorization for housekeeping genes after Hurst and colleagues (22) is shown in Supplementary Figure S2. (D) The NSL complex prevalently binds to dispersed promoters ('broad with peak' and 'weak peak' promoters) over peaked promoters (Welch two sample t -test, $P < 2.2 \times 10^{-16}$). Density plot of NSL1 binding at genes, which were grouped according their transcriptional start site usage in 'peaked' promoters, 'broad with peaked' promoters and 'weak peak' promoters using the data of (19). The window is split for NSL1-bound (right) and -unbound (left) promoters.

displaying gene activity as a function of polymerase promoter binding (Figure 1B, right) or Affymetrix RNA expression profiling (data not shown), in agreement with previous studies examining other markers of the NSL complex (11,12).

The NSL complex specifically binds promoters of most housekeeping genes

As noted above, the NSL1 profile is very similar in nuclei of different sex and developmental stage despite significant expression differences (Supplementary Figure S2A and D). This indicates that the NSL complex may associate with 'housekeeping' genes, which are equally expressed in these diverse tissues. To test this hypothesis, we classified genes as 'housekeeping' or 'differentially regulated' according to their expression variation index, i.e. the standard deviation of expression, when compared between several *Drosophila* tissues (21). According to this classification the NSL complex showed a significant preference for 'housekeeping' over 'differentially regulated' genes (Figure 1C). The same conclusion was reached when 'housekeeping' genes were classified according to the more exclusive definition of Hurst and colleagues (22) (Supplementary Figure S2E). This conclusion is further illustrated by a gene ontology (GO) analysis of bound and unbound genes, which revealed that active NSL-bound genes are enriched in housekeeping functions such as 'cofactor biosynthetic processes', 'microtubule-based processes', 'protein complex biogenesis' (Supplementary Figure S3), whereas active genes which are not bound by the NSL complex are enriched in categories such as 'sensory perception', 'cell adhesion' and 'tissue developmental genes' (Supplementary Figure S4).

Recent improvements in high-throughput RNA profiling techniques facilitated quantitative mapping of TSSs at base pair resolution (19,20). Whereas some promoters possess well-defined TSS, where transcription reliably initiates within a few base pairs ('focused' or 'peaked' promoters), many promoters show a dispersed zone of

transcription initiation of up to a few hundred base pairs, which may be dominated by a major TSS ('broad promoters') or not ('weak peak promoter') (20). Notably, differentially regulated genes tend to have peaked promoters whereas housekeeping genes are enriched for broad or weak promoters (19). Concordantly, we found that the NSL complex is strongly overrepresented at promoters of the latter classes (Figure 1D).

The NSL complex activates only a specific subset of bound genes

It has remained controversial whether NSL target genes are activated or repressed after RNAi ablation of NSL complex components (11–13). Akthar and coworkers observed that similar fractions of NSL target genes were up- or downregulated following RNAi against MOF, NSL3 and MBD-R2 and subsequent microarray-based transcriptome profiling (11,13). By contrast, we found that the transcription of genes that had the NSL subunit MBD-R2 bound was mostly reduced when MBD-R2 levels were lowered (12). However, since MBD-R2 is the only NSL complex subunit that was suggested to interact with components of the DCC (10), it was necessary to exclude indirect effects. We therefore examined the expression of NSL target genes after depletion of the core subunit of the NSL complex, NSL1.

RNAi against NSL1 in S2 cells efficiently depleted the protein as examined by immunoblotting and IFM (Supplementary Figure S1). Genome-wide transcriptome profiling of the NSL1-depleted cells led to the down-regulation of a considerable fraction of genes (Figure 2A), most of which had been classified as 'NSL-bound' before (Figure 2B). This is consistent with reporter gene assays where the transcription brought about by tethering MOF to a model promoter was diminished upon NSL1 depletion (for details, see Supplementary Figure S5). Importantly, the expression of the majority of NSL1 target genes was unchanged (Figure 2A), such that only 20–30% of them (depending on the threshold) required

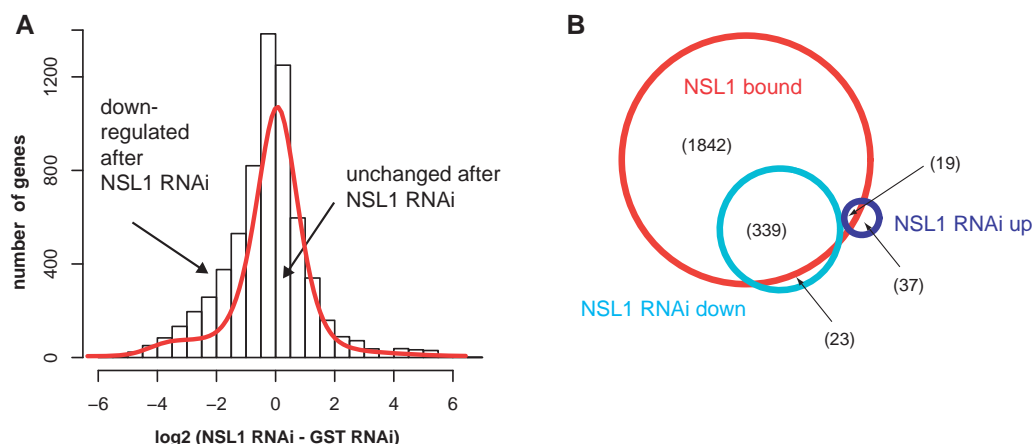


Figure 2. The NSL complex activates only a subset of its target promoters. (A) The histogram depicts transcriptome changes upon NSL1 depletion in S2 cells. The overlaid modeled normal distribution (red) reveals a skew toward the population of down-regulated genes. The gene set was filtered for active genes based on the transcriptome of control cells. (B) Proportional Venn diagram depicts genes 2-fold up- or down-regulated after RNAi against NSL1, respectively, and NSL1-bound genes (NSL1 ChIP-chip in S2 cells). Numbers in parentheses indicate the size of the respective genes sets. Only active genes (total polymerase promoter occupancy determined by the modENCODE consortium) represented on our custom-tailored microarray tiling array are shown.

NSL1 for proper expression. The same trend had been observed earlier in the context with MBD-R2 (12) (and data not shown). The MBD-R2 ChIP-chip profile and the MBD-R2 RNAi transcriptome data are indeed very similar to the NSL1 data (Supplementary Figures S2C and S5C), arguing that they form a functional complex bound to chromatin.

We next asked whether the genes that were activated by the NSL complex coded for related housekeeping functions. The GO classification revealed that the genes whose expression was diminished upon NSL1 depletion were enriched in genes involved in nucleic acid metabolism, such as genes involved in transcription, RNA processing, translation, DNA replication and DNA repair (Supplementary Figure S6). Evidently, the NSL complex only activates a specific subset of the many housekeeping genes. In order to explore whether the promoters of NSL-responsive genes could be recognized by a combination of *cis*-elements and *trans*-factors, we set out to identify chromatin proteins with genome binding profiles related to the NSL complex and to investigate whether the promoters regulated by NSL shared particular core promoter motifs.

The NSL complex co-occupies target promoters together with the chromatin remodeler NURF and the histone methyltransferase Trithorax

The NSL1 ChIP-chip profile in S2 cells allowed a direct comparison with the chromatin profiles recorded by the modENCODE consortium (17), which used the same cell line and the same profiling technique. We mined the modENCODE data for profiles of general chromatin factors (excluding sequence-specific transcription factors) and histone modifications, which are preferentially enriched at promoters (see 'Materials and Methods' section for a detailed discussion on selection algorithm). We created a pairwise correlation matrix for 23 selected protein and histone modification profiles and performed an unsupervised hierarchical clustering to reveal the extent of correlation with the NSL complex. We found the profiles of the interband protein Chromator, the WD40-repeat protein WDS, the NURF complex subunit NURF301 and the methyltransferase Trithorax highly correlated with the NSL complex profile (Figure 3A and B; Supplementary Figure S7). Chromator had been found in an early NSL complex purification (10) but could not be recovered in more recent experiments (11,12), possibly due to more transient or indirect interaction. Notably, 5–15% of promoters which contain NSL1, MBD-R2, WDS, NURF301 and Trithorax lack Chromator. The WD40-repeat protein WDS consistently copurifies with NSL complex members (10–12) and other chromatin complexes including the *Drosophila* ATAC acetyltransferase complex (26) and mammalian MLL methyltransferase complexes (27,28). NURF301 is the diagnostic marker subunit of the Imitation Switch (ISWI)-containing nucleosome remodeling factor NURF, which stimulates transcription by remodeling promoter nucleosomes (29,30). Trithorax was originally described to counteract the repression of homeotic genes

by the polycomb group proteins (31–33). More recently, genome-wide ChIP-chip studies have indicated a widespread binding of Trithorax to many promoters (34,35).

The pairwise relationships between the tested factors are further illustrated by the scatter plots depicted in Figure 3C, which emphasize that the NSL complex, WDS, Chromator, NURF301 and Trithorax co-occupy target promoters at linearly proportional levels (Figure 3C). Promoters which are strongly bound by the NSL complex are also highly enriched for NURF301, Chromator and Trithorax. The same strong correlation can be seen in an unbiased analysis using all microarray probe signals, confirming the promoter-focused analysis described above (Supplementary Figure S7B).

Searching for factors enriched at promoters we found the heterochromatin protein 1c (HP1c) and, consistent with previous results (36), the insulator proteins BEAF32 and CP190 (37) enriched at housekeeping promoters (Supplementary Figure S8). These factors localize to minor subsets of the NSL/Chromator/NURF301/Trithorax target promoters (Figures 3 and 5; Supplementary Figures S7B and 11). Importantly, the presence of BEAF32, CP190 and HP1c determines whether the bound NSL complex functions as an activator or not (see below).

Quantitative NSL1 binding correlates best with the DNA replication-related element

Conceivably, the association of the NSL complex and its colocalized chromatin modifiers may be determined by a particular core promoter architecture. Different promoters are characterized by the presence and combination of a range of sequence motifs that provide contact surfaces for general transcription factors and, therefore, modulate the formation of the transcription pre-initiation complex (38–40). The core promoter sequence motifs can be classified as canonical core promoter motifs which have fixed positions with regard to the TSS, such as the TATA box, the MTE (motif ten element), the DPE (downstream core promoter element) and the INR (initiator), or as motifs with weaker positional information (Ohler 1, Ohler 5, Ohler 6, Ohler 7, Ohler 8 and DRE) (25,41). Canonical core promoter motifs are enriched in peaked promoters, whereas weakly positioned motifs are characteristic of dispersed promoters. The mechanisms of action of most dispersed elements are unknown [with the exception of the DRE (39)].

Since NSL1 peaks within the core promoters of genes with dispersed transcriptional start sites (Figure 1A and D) we investigated whether the NSL complex is associated with a specific set of core promoter motifs. We first characterized the core promoter motifs with regard to their distribution at active housekeeping and differentially regulated genes (Supplementary Figure S9). As the motifs deviate from their defined consensus sequences in many cases, a similarity score (motif score) was calculated for each promoter reflecting the similarity of the sequence to any of the ten promoter consensus motifs described by Ohler and colleagues (25) (see 'Materials and Methods' section). We found that over 70% of all active promoters

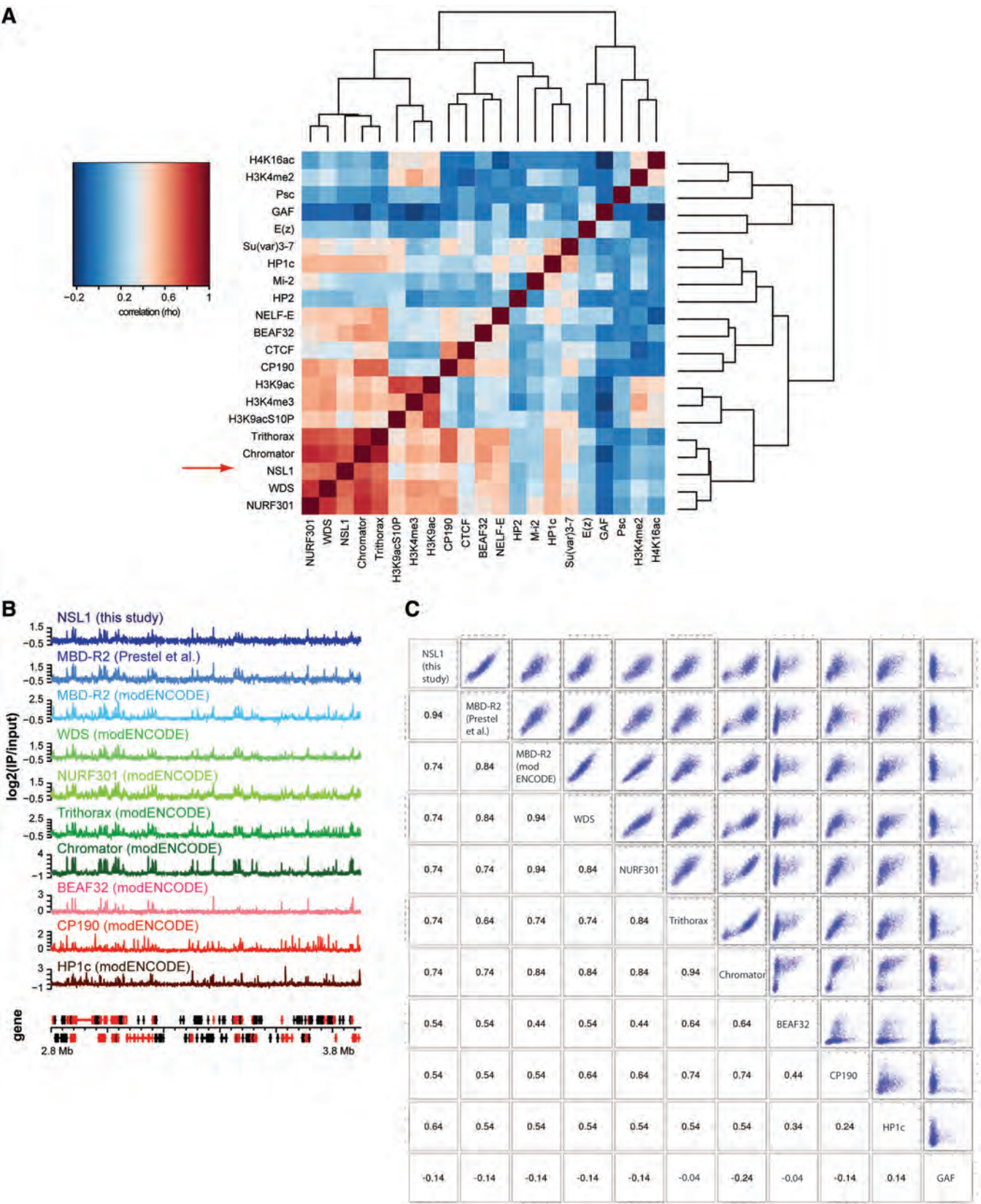


Figure 3. The NSL complex cooccupies target promoters together with the chromatin remodeler NURF and the histone methyltransferase Trithorax. (A) Heat map visualized correlation matrix of promoter-enriched chromatin factors and histone modifications at active genes. Pairwise Spearman correlations were calculated using only active autosomal genes. The dendrogram shows the hierarchical clustering after which the matrix was sorted. (B) ChIP-chip profile of the indicated proteins along a representative region of the chromosome arm 2L. The gene structure is indicated below (active genes are red). (C) Pairwise scatter plot of promoter binding for each indicated factor using only active autosomal genes. Spearman non-parametric correlation coefficients are provided for each pair.

can be described by these ten motifs, indicating that our analysis is representative (Supplementary Figure S9). In agreement with previous analyses (41–43), the promoter motifs INR, MTE and DPE were clearly overrepresented in differentially regulated genes, which fits their enrichment at peaked promoters (Supplementary Figures S8 and S9). Accordingly, housekeeping genes are enriched for the motifs DRE, Ohler1, Ohler 5, Ohler 6 and Ohler 7.

Being able to characterize the core promoter motifs allowed us to examine whether there is a differential association of NSL with any of them. Active genes were categorized either as NSL targets or as non-targets based on their NSL complex promoter occupancy. As expected, the NSL1 target genes are enriched for the housekeeping promoter motifs DRE, Ohler 1, Ohler 5, Ohler 6 and Ohler 7 and depleted for TATA, INR, MTE and DPE. Consistently, when we performed *de novo* motif analysis of the sequences covered by the NSL1 ChIP-Seq peaks (11), we again obtained the same motifs (Supplementary Figure S10). This confirms that the NSL1 peaks at core promoter motifs are diagnostic for housekeeping genes.

Is there any correlation between the ‘strength’ of NSL1 binding and how well an underlying motif matches its consensus sequence? In order to address this question we used the NSL1 ChIP-Seq data set (11), which due to its good dynamic range allowed to categorize the ChIP-Seq peak score as a surrogate for binding ‘strength’. We binned the ChIP-Seq peaks in equally sized groups according to their peak score [determined by SPP, (24)] and displayed the fraction of promoters bound by a specific group at a given motif score (Figure 4A, left). Among the ten tested core promoter motifs the DRE motif, and to a lesser extent motif Ohler 7, are the only motifs with scores that correlate with the NSL1 ChIP-Seq peak score. This suggests that DRE-containing promoters (and those containing the less abundant Ohler 7 motif) primarily contain NSL complex targeting clues (Figure 4B).

The combination of chromatin factors and core promoter motifs enhance the prediction of NSL-regulated promoters

Whether or not a promoter-bound transcription factor engages in active regulation often depends on the context of close-by *cis* elements and interacting factors (44). This appears to be the case for the NSL complex, as we showed that the complex only activates a subset of the promoters it associates with. NSL binds with high preference to a set of housekeeping promoter motifs and its binding ‘strength’ correlates best with the presence of the DRE motif. Can the subset of these NSL targets whose transcription is diminished after depletion of NSL (i.e. those promoters at which the complex is functional as an activator) be distinguished at the sequence level? We grouped active genes according to their core promoter motif class (see ‘Materials and Methods’ section) and monitored the transcriptome changes after NSL depletion for each group. Strikingly, only promoters containing the core promoter motif ‘Ohler 5’ were strongly enriched for NSL complex functional sites (Figure 5A). We note that ‘Ohler 5’-containing promoters do not show the strongest correlation to NSL binding strength (Figure 4B)

suggesting that quantitative differences in factor binding are not directly translated into a functional output.

We had observed that most promoters bound by HP1c, BEAF32 and CP190 are among those also occupied by the NSL complex (Figure 3 and Supplementary Figure S11). Most of the HP1c, BEAF32 and CP190 binding occur at distinct subsets as the three factors only colocalize at a minority of sites (Supplementary Figure S11). Intriguingly, promoters bound by any of the three factors HP1c, BEAF32 or CP190 are obviously underrepresented among the genes, whose transcription is activated by the NSL complex (Figure 5B and Supplementary Figure S11).

In summary, the data suggest that the functionality of a promoter-associated NSL complex is modulated by positive effectors (e.g. unidentified interactors of the ‘Ohler 5’ element) and negative regulators (HP1c and the insulator proteins BEAF32 and CP190).

DISCUSSION

In this study, we show that the NSL complex is a potential coactivator, which binds to many active genes, but regulates only a specific subset of them. In our efforts to describe the circumstances that define complex association and function, we considered the contributions of two major parameters: the diverse DNA sequences around the core promoters, which are characterized by combinations of recurring sequence motifs, and the association of chromatin regulators that have recently been mapped by the modENCODE consortium. Combining these diverse data sets, we were able to improve the prediction toward whether the transcription of an NSL-bound gene is modulated by the NSL complex. To our knowledge, this is the first systematic study demonstrating the usefulness of this type of data integration.

The NSL complex is a transcription cofactor dedicated to housekeeping genes

Following our observation that the NSL complex binds to only a subset of all active promoters, we discovered that the target genes were mostly housekeeping genes. This was surprising as to our knowledge so far no transcription coregulator dedicated to housekeeping genes is known. This may simply reflect the fact that historically the mechanisms underlying differential transcription regulation received more attention. Several lines of evidence support the conclusion that the NSL complex preferentially localizes to the majority of housekeeping promoters. (i) We do not detect significant differences in the global chromatin binding profile of NSL complex members in cells of different sex or developmental stage. (ii) Genes that have NSL bound at their promoters show little expression variation among different tissues as compared to active genes that lack the NSL complex. (iii) NSL-bound promoters are depleted of sequence motifs known to be enriched in genes differentially regulated during development and in tissue homeostasis (38). (iv) GO analysis of the active NSL-bound genes revealed an overrepresentation of categories for housekeeping functions, whereas the converse data set of active genes not bound by NSL

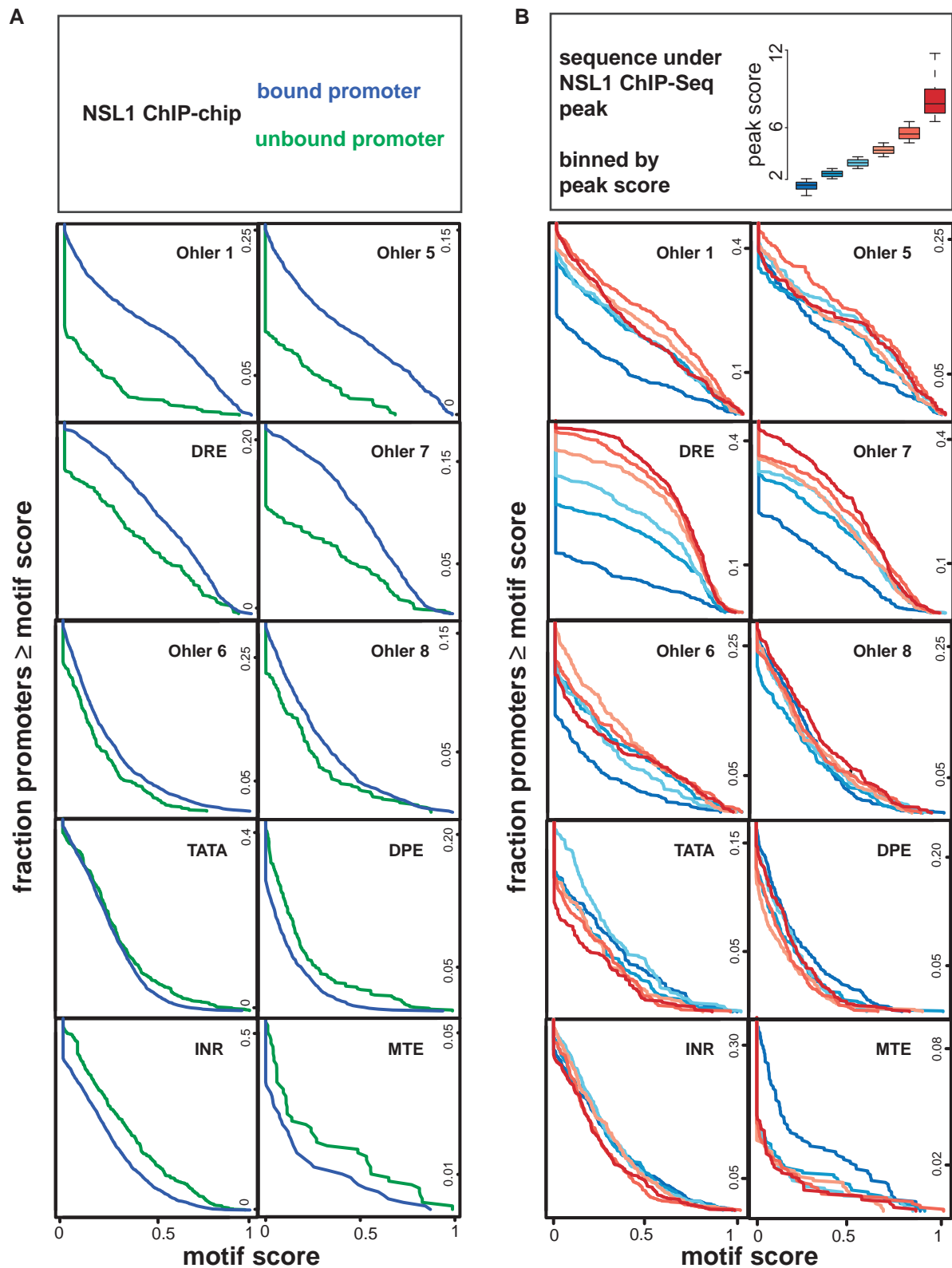


Figure 4. NSL1 binding correlates best with the DRE motif score. (A) Promoters containing motifs Ohler 1, Ohler 5, DRE and Ohler 7 are overrepresented among the NSL complex target genes. Active genes were grouped into NSL1-bound and NSL1-unbound genes, as determined by NSL1 ChIP-chip in S2 cells. The curves display the cumulative fraction of promoters, which exceed a given motif score. The motif score is derived from the information content of the position weight matrix, as taken from Ohler and colleagues (25), see 'Materials and Methods' section for details). (B) NSL1 binding 'strength' linearly scales to the DRE motif score. NSL1 ChIP-Seq peaks (11) are grouped in equal-sized bins according to their peak score. Similar to (A), the fraction of promoters is displayed relative to the motif score.

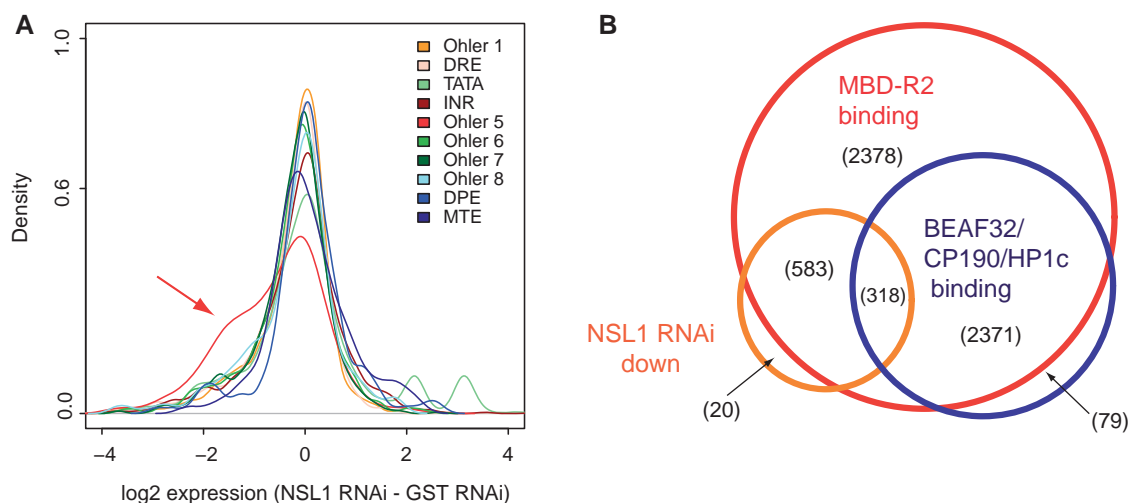


Figure 5. A combination of chromatin factors and core promoter motifs enhance the prediction of NSL transcriptional targets. (A) The core promoter motif Ohler 5 is overrepresented in genes, which are regulated by the NSL complex ($P < 2.2 \times 10^{-16}$, one-sample *t*-test). The density plot of the transcriptome changes after RNAi against NSL1 for groups of genes enriched for the ten core promoter motifs defined by Ohler *et al.* (25). In the presented analysis only active NSL-bound genes were considered (using the modENCODE MBD-R2 data set (17), which allows the comparison to a larger transcriptome data set). The same conclusion was reached if all active genes (without filtering for NSL complex binding) or different 'NSL complex binding' criteria (such as NSL1 binding) were employed (data not shown). (B) Venn diagram representation of gene sets bound by the NSL complex (modENCODE MBD-R2 ChIP-chip profile), down-regulated after ablation of the NSL complex and bound by at least one of the three investigated proteins BEAF32, CP190 and HP1c (17). Numbers in brackets represent the number of genes in the respective group. The modENCODE MBD-R2 data set served as a surrogate for the NSL complex to minimize technical variation in the comparison with the BEAF32/CP190/HP1c data sets. Similar results were obtained using the NSL1 data set and the MBD-R2 (12) data set, respectively.

present diverse categories including 'developmental programs' and 'acute signaling'. Other chromatin constituents, like HP1c and the insulator proteins BEAF32 and CP190 also interact preferentially with housekeeping gene promoters, as previously shown by Ohler and colleagues (36), but these factors bind to a much more limited number of genes in this class. Our analysis supports the concept of global coregulation of functionally related gene classes by common cofactors.

The developmental regulators NURF and Trithorax colocalize with NSL at housekeeping promoters

The extensive colocalization of the NSL complex with the methyltransferase Trithorax and the chromatin remodeler NURF is puzzling since those factors are best known as regulators of transcription of very restricted sets of genes (developmental and highly inducible genes) (30,32), and only recently has their extensive genome-wide localization at many active gene promoters been noticed (34,35,45,46). Conceivably, these three complexes cooperate to regulate the transcription of housekeeping genes at the level of chromatin organization and/or transcription initiation. This hypothesis is supported by previous reports of biochemical or genetic interactions between components of the three factors. A genetic interaction between the *Xenopus* BPTF (the NURF301 homolog) and *Xenopus* WDR5 (a homolog of the NSL subunit WDS) has been reported (47). Furthermore, Dou *et al.* (27) described a 'supercomplex' containing the human NSL as well as the MLL1 complexes [MLL1 is homologous to *Drosophila* Trithorax].

At present it is not clear whether NURF and Trithorax-containing complexes contribute to the targeting of the NSL complex (or vice versa), or whether all three regulators are attracted by an additional common denominator of target promoters. None of the three complexes contains any specific DNA-binding subunit. NURF can be recruited to inducible genes via direct interactions between the large NURF301 subunit and transcription factors, such as the GAGA factor (29) or the ecdysone receptor (48). However, these interactions certainly do not explain the widespread targeting of NURF to housekeeping genes *in vivo* reported here. We noted a good quantitative correlation between the NSL1 binding levels and the DRE core promoter motif score, which opens the possibility that a DRE-recognizing factor may stimulate NSL recruitment. One candidate for such a factor is DREF, which has been isolated as a DRE binding factor (49). DREF may also contribute to the recruitment of NURF, since an association of DREF with NURF has been observed in a much larger complex based on the immunoaffinity purification of the TATA box binding protein (TBP)-related factor TRF2 (39).

In addition to direct recruitment by DNA-binding proteins, transcription cofactors may be tethered by specific local histone modifications through recognition domains (50). It is likely that this principle will also contribute to the observed colocalization of NSL, NURF and Trithorax complexes. Trithorax (the *Drosophila* MLL1 homolog) is an enzyme that methylates histone H3 at lysine 4 (H3K4me3), a mark that characterizes active promoters (46). Interestingly, WDS, which copurifies with NSL complexes from *Drosophila* and mammalian cells

(27) has been shown to preferentially interact with methylated H3K4 (28). The mammalian homolog of NURF301 (BPTF) also recognizes mononucleosomes marked with methylated H3K4 and acetylated H4K16ac through its PHD finger and bromodomain, respectively (51). Acetylation of H4K16 by MOF in the NSL complex may, therefore, contribute to the local enrichment of NURF at target promoters. Our study gives rise to numerous testable hypotheses as to the nature of the interaction network that leads to the observed selective targeting of the NSL complex.

The NSL complex only regulates a subset of target promoters

The detailed analysis of the transcriptional effects of the NSL complex revealed that the NSL complex regulates only a subset of bound genes. Such a situation is not without precedent as it has been shown for a number of transcription factors that many binding events appear to be non-functional (44). In fact, it is a major challenge to predict the functional sites from the interaction profiles of single factors as functionality is frequently determined by the local clustering of binding sites, synergism between colocalized proteins, and recently, chromatin accessibility (52,53). Accordingly, we favor the idea that a combination of chromatin factors and core promoter elements determines the activity of the NSL complex at any target promoter. An even more immediate influence of promoter DNA on interacting proteins may be imagined as a direct effect of a DNA sequence on the conformation and, therefore, the activity of a bound transcription factor has been described (54).

Alternatively, it is possible that the default state of every chromatin-bound NSL complex is functional, but that the realization of this potential is restricted by negative factors. We found that the presence of either one of the three proteins HP1c, BEAF32 or CP190 correlated with lack of NSL1 regulation. Insulator binding proteins like BEAF32 and CP190 are known to decrease enhancer-promoter interactions, which may lead to decreased transcriptional output. Interestingly, antagonistic roles for BEAF32 and DREF have been suggested for some overlapping *in vivo* binding sites (55). Resolving the mechanistic intricacies of complex promoter regulation remains a challenging task for future endeavors.

ACCESSION NUMBER

Microarray accession number: GSE30991.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–11 and Supplementary References [57].

ACKNOWLEDGMENTS

We are grateful to A. Mitterweger for help with ChIPs in S2 cells, H. Mitlöhner for assistance in fly maintenance,

H. Saumweber for providing the lamin antibody, D. Martin and K. Maier of the Gene Center Affymetrix Microarray Platform for help with microarray experiments and A. Tresch and members of the Becker lab for helpful discussions. C.F. carried out the ChIP-chip and gene expression profiling experiments and analyzed the ChIP-chip, ChIP-Seq and gene expression data. M.P. developed and characterized the NSL1 antibody. M.P. and C.F. performed the RNAi, immunoblot, IFM and reporter gene ChIP experiments, and M.P. performed the luciferase reporter gene experiments. H.H. and J.S. performed the core promoter analysis. T.S. supervised C.F. in the bioinformatic analyses. P.B.B. and C.F. conceived the study and wrote the article. All authors read and approved the article.

FUNDING

Deutsche Forschungsgemeinschaft through SFB-TR5 (to P.B.B.) and SFB646 (to J.S.); the Gottfried-Wilhelm-Leibniz Programme (to P.B.B.); the International Max-Planck Research School in Munich (to C.F.). Funding for open access charge: Grant from the Deutsche Forschungsgemeinschaft (to P.B.B.).

Conflict of interest statement. None declared.

REFERENCES

- Clapier, C.R. and Cairns, B.R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.*, **78**, 273–304.
- Kwon, S.Y., Xiao, H., Wu, C. and Badenhorst, P. (2009) Alternative splicing of NURF301 generates distinct NURF chromatin remodeling complexes with altered modified histone binding specificities. *PLoS Genet.*, **5**, e1000574.
- Tsukiyama, T. and Wu, C. (1995) Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell*, **83**, 1011–1020.
- Smith, E.R., Cayrou, C., Huang, R., Lane, W.S., Cote, J. and Lucchesi, J.C. (2005) A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Mol. Cell. Biol.*, **25**, 9175–9188.
- Taipale, M., Rea, S., Richter, K., Vilar, A., Lichter, P., Imhof, A. and Akhtar, A. (2005) hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. *Mol. Cell. Biol.*, **25**, 6798–6810.
- Straub, T. and Becker, P.B. (2007) Dosage compensation: the beginning and end of generalization. *Nat. Rev.*, **8**, 47–57.
- Gelbart, M.E., Larschan, E., Peng, S., Park, P.J. and Kuroda, M.I. (2009) Drosophila MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. *Nat. Struct. Mol. Biol.*, **16**, 825–832.
- Larschan, E., Bishop, E.P., Kharchenko, P.V., Core, L.J., Lis, J.T., Park, P.J. and Kuroda, M.I. (2011) X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. *Nature*, **471**, 115–118.
- Prestel, M., Feller, C. and Becker, P.B. (2010) Dosage compensation and the global re-balancing of aneuploid genomes. *Genome Biol.*, **11**, 216.
- Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J. et al. (2006) Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. *Mol. Cell*, **21**, 811–823.
- Raja, S.J., Charapitsa, I., Conrad, T., Vaquerizas, J.M., Gebhardt, P., Holz, H., Kadlec, J., Fraterman, S., Luscombe, N.M. and Akhtar, A.

- (2010) The nonspecific lethal complex is a transcriptional regulator in *Drosophila*. *Mol. Cell*, **38**, 827–841.
12. Prestel, M., Feller, C., Straub, T., Mitlohner, H. and Becker, P.B. (2010) The activation potential of MOF is constrained for dosage compensation. *Mol. Cell*, **38**, 815–826.
 13. Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P. and Akhtar, A. (2008) Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in *Drosophila*. *Cell*, **133**, 813–828.
 14. Straub, T., Neumann, M.F., Prestel, M., Kremmer, E., Kaether, C., Haass, C. and Becker, P.B. (2005) Stable chromosomal association of MSL2 defines a dosage-compensated nuclear compartment. *Chromosoma*, **114**, 352–364.
 15. Gilfillan, G.D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B. and Becker, P.B. (2006) Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex. *Genes Dev.*, **20**, 858–870.
 16. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A. and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, **18**(Suppl. 1), S96–S104.
 17. Celniker, S.E., Dillon, L.A., Gerstein, M.B., Gunsalus, K.C., Henikoff, S., Karpen, G.H., Kellis, M., Lai, E.C., Lieb, J.D., MacAlpine, D.M. *et al.* (2009) Unlocking the secrets of the genome. *Nature*, **459**, 927–930.
 18. Regnard, C., Straub, T., Mitterweger, A., Dahlsveen, I.K., Fabian, V. and Becker, P.B. (2011) Global analysis of the relationship between JIL-1 kinase and transcription. *PLoS Genet.*, **7**, e1001327.
 19. Hoskins, R.A., Landolin, J.M., Brown, J.B., Sandler, J.E., Takahashi, H., Lassmann, T., Yu, C., Booth, B.W., Zhang, D., Wan, K.H. *et al.* (2011) Genome-wide analysis of promoter architecture in *Drosophila melanogaster*. *Genome Res.*, **21**, 182–192.
 20. Ni, T., Corcoran, D.L., Rach, E.A., Song, S., Spana, E.P., Gao, Y., Ohler, U. and Zhu, J. (2010) A paired-end sequencing strategy to map the complex landscape of transcription initiation. *Nat. Methods*, **7**, 521–527.
 21. Chintapalli, V.R., Wang, J. and Dow, J.A. (2007) Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat. Genet.*, **39**, 715–720.
 22. Weber, C.C. and Hurst, L.D. (2011) Support for multiple classes of local expression clusters in *Drosophila melanogaster*, but no evidence for gene order conservation. *Genome Biol.*, **12**, R23.
 23. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
 24. Kharchenko, P.V., Tolstorukov, M.Y. and Park, P.J. (2008) Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat. Biotechnol.*, **26**, 1351–1359.
 25. Ohler, U., Liao, G.C., Niemann, H. and Rubin, G.M. (2002) Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol.*, **3**, RESEARCH0087.
 26. Suganuma, T., Gutierrez, J.L., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M. and Workman, J.L. (2008) ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nat. Struct. Mol. Biol.*, **15**, 364–372.
 27. Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L. and Roeder, R.G. (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell*, **121**, 873–885.
 28. Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H. and Allis, C.D. (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell*, **121**, 859–872.
 29. Xiao, H., Sandaltzopoulos, R., Wang, H.M., Hamiche, A., Ranallo, R., Lee, K.M., Fu, D. and Wu, C. (2001) Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. *Mol. Cell*, **8**, 531–543.
 30. Badenhorst, P., Voas, M., Rebay, I. and Wu, C. (2002) Biological functions of the ISWI chromatin remodeling complex NURF. *Genes Dev.*, **16**, 3186–3198.
 31. Simon, J.A. and Tamkun, J.W. (2002) Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. *Curr. Opin. Genet. Dev.*, **12**, 210–218.
 32. Ringrose, L. and Paro, R. (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.*, **38**, 413–443.
 33. Klymenko, T. and Muller, J. (2004) The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Reports*, **5**, 373–377.
 34. Schuettengruber, B., Ganapathi, M., Leblanc, B., Portoso, M., Jaschek, R., Tolhuis, B., van Lohuizen, M., Tanay, A. and Cavalli, G. (2009) Functional anatomy of polycomb and trithorax chromatin landscapes in *Drosophila* embryos. *PLoS Biol.*, **7**, e13.
 35. Enderle, D., Beisel, C., Stadler, M.B., Gerstung, M., Athri, P. and Paro, R. (2011) Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. *Genome Res.*, **21**, 216–226.
 36. Rach, E.A., Winter, D.R., Benjamin, A.M., Corcoran, D.L., Ni, T., Zhu, J. and Ohler, U. (2011) Transcription initiation patterns indicate divergent strategies for gene regulation at the chromatin level. *PLoS Genet.*, **7**, e1001274.
 37. Bushey, A.M., Dorman, E.R. and Corces, V.G. (2008) Chromatin insulators: regulatory mechanisms and epigenetic inheritance. *Mol. Cell*, **32**, 1–9.
 38. Juven-Gershon, T. and Kadonaga, J.T. (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol.*, **339**, 225–229.
 39. Hochheimer, A., Zhou, S., Zheng, S., Holmes, M.C. and Tjian, R. (2002) TRF2 associates with DREF and directs promoter-selective gene expression in *Drosophila*. *Nature*, **420**, 439–445.
 40. Goodrich, J.A. and Tjian, R. (2010) Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. *Nat. Rev.*, **11**, 549–558.
 41. Ohler, U. (2006) Identification of core promoter modules in *Drosophila* and their application in accurate transcription start site prediction. *Nucleic Acids Res.*, **34**, 5943–5950.
 42. FitzGerald, P.C., Sturgill, D., Shyakhtenko, A., Oliver, B. and Vinson, C. (2006) Comparative genomics of *Drosophila* and human core promoters. *Genome Biol.*, **7**, R53.
 43. Rach, E.A., Yuan, H.Y., Majoros, W.H., Tomancak, P. and Ohler, U. (2009) Motif composition, conservation and condition-specificity of single and alternative transcription start sites in the *Drosophila* genome. *Genome Biol.*, **10**, R73.
 44. MacArthur, S., Li, X.Y., Li, J., Brown, J.B., Chu, H.C., Zeng, L., Grondana, B.P., Hechmer, A., Simirenko, L., Keranen, S.V. *et al.* (2009) Developmental roles of 21 *Drosophila* transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. *Genome Biol.*, **10**, R80.
 45. Schwartz, Y.B., Kahn, T.G., Stenberg, P., Ohno, K., Bourgon, R. and Pirrotta, V. (2010) Alternative epigenetic chromatin states of polycomb target genes. *PLoS Genet.*, **6**, e1000805.
 46. Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T. *et al.* (2011) Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature*, **471**, 480–485.
 47. Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P. *et al.* (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature*, **442**, 86–90.
 48. Badenhorst, P., Xiao, H., Cherbas, L., Kwon, S.Y., Voas, M., Rebay, I., Cherbas, P. and Wu, C. (2005) The *Drosophila* nucleosome remodeling factor NURF is required for Ecdysteroid signaling and metamorphosis. *Genes Dev.*, **19**, 2540–2545.
 49. Hirose, F., Yamaguchi, M., Handa, H., Inomata, Y. and Matsukage, A. (1993) Novel 8-base pair sequence (*Drosophila* DNA replication-related element) and specific binding factor involved in the expression of *Drosophila* genes for DNA polymerase alpha and proliferating cell nuclear antigen. *J. Biol. Chem.*, **268**, 2092–2099.

50. Ruthenburg, A.J., Li, H., Patel, D.J. and Allis, C.D. (2007) Multivalent engagement of chromatin modifications by linked binding modules. *Nat. Rev. Mol. Cell Biol.*, **8**, 983–994.
51. Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T.W., Patel, D.J. *et al.* (2011) Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell*, **145**, 692–706.
52. Li, X.Y., Thomas, S., Sabo, P.J., Eisen, M.B., Stamatoyannopoulos, J.A. and Biggin, M.D. (2011) The role of chromatin accessibility in directing the widespread, overlapping patterns of *Drosophila* transcription factor binding. *Genome Biol.*, **12**, R34.
53. Kaplan, T., Li, X.Y., Sabo, P.J., Thomas, S., Stamatoyannopoulos, J.A., Biggin, M.D. and Eisen, M.B. (2011) Quantitative models of the mechanisms that control genome-wide patterns of transcription factor binding during early *Drosophila* development. *PLoS Genet.*, **7**, e1001290.
54. Meijssing, S.H., Pufall, M.A., So, A.Y., Bates, D.L., Chen, L. and Yamamoto, K.R. (2009) DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science*, **324**, 407–410.
55. Hart, C.M., Cuvier, O. and Laemmli, U.K. (1999) Evidence for an antagonistic relationship between the boundary element-associated factor BEAF and the transcription factor DREF. *Chromosoma*, **108**, 375–383.
56. Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y. and Adelman, K. (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science*, **327**, 335–338.
57. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.

3.3.3 Supplementary data and figures

Supplementary Data

The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset

Christian Feller¹, Matthias Prestel¹, Holger Hartmann², Tobias Straub¹, Johannes Söding²
and Peter B. Becker¹

¹Adolf-Butenandt-Institute and ²Gene Center / Department of Biochemistry, Center for Integrated Protein Science Munich of the Ludwig-Maximilians-University, Munich, Germany

Corresponding author

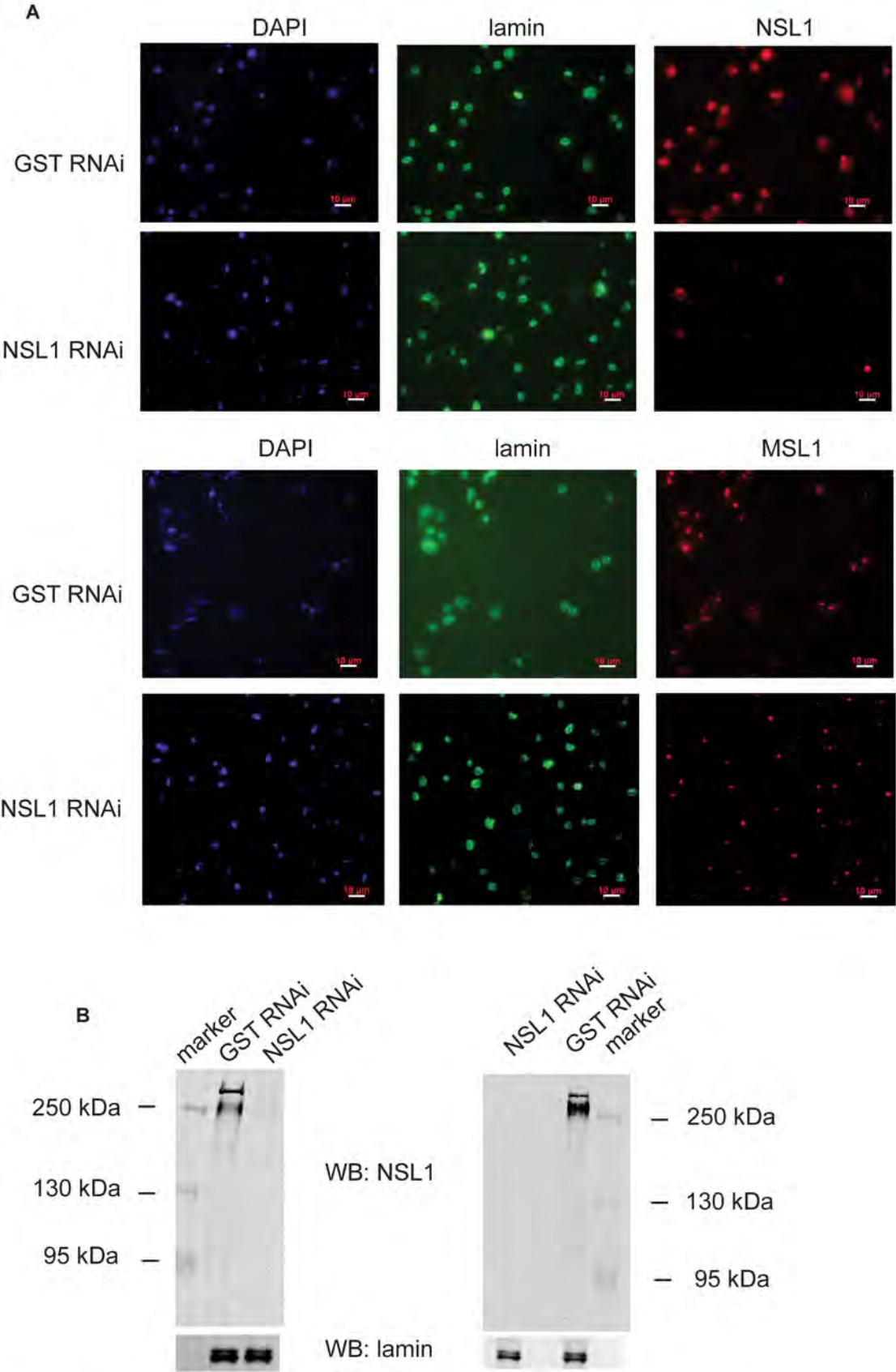
Peter B. Becker

Adolf-Butenandt-Institute,
Schillerstrasse 44, 80336 München, Germany

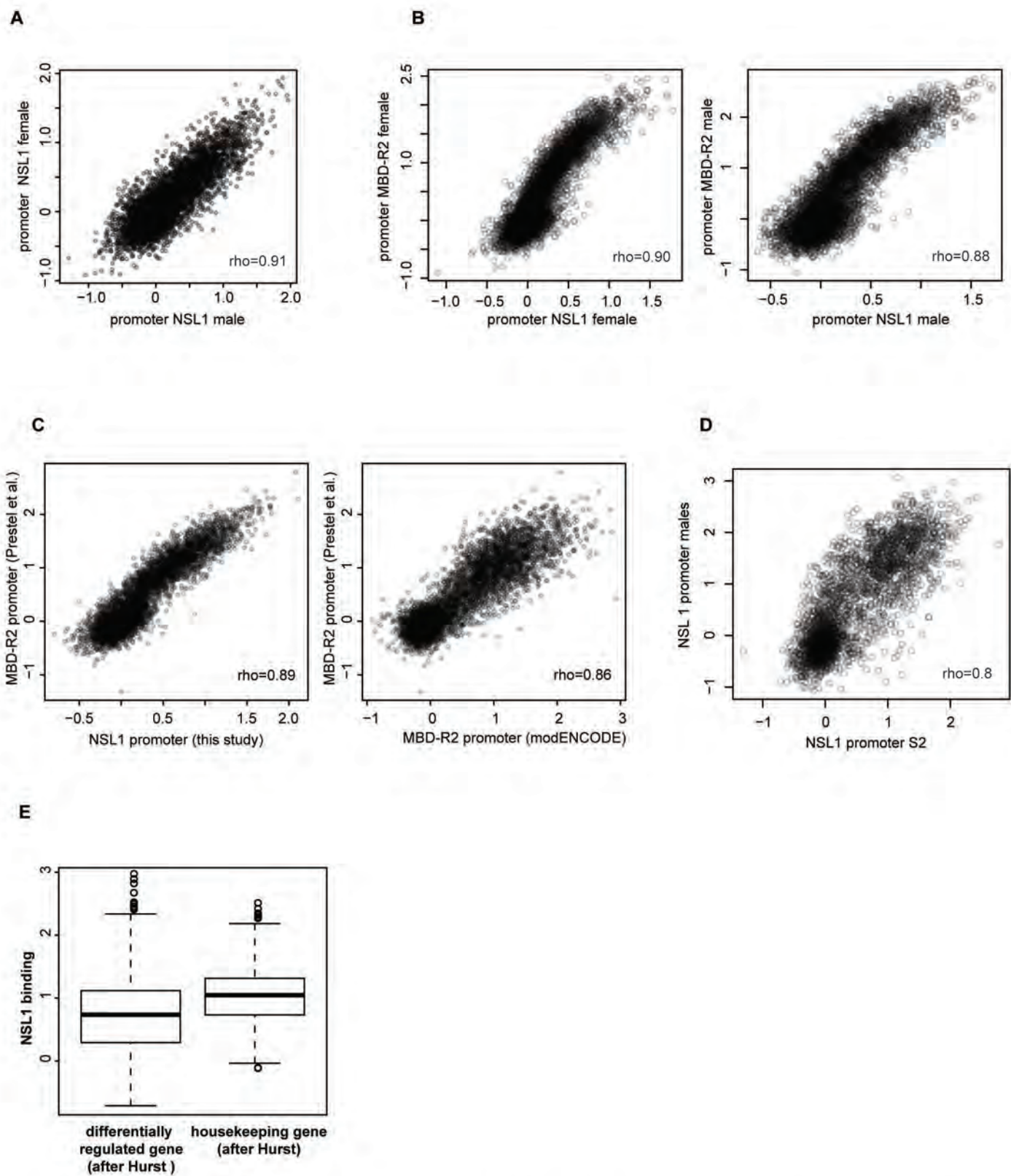
Phone: 089-2180-75-427

Fax: 089-2180-75-425

Email: pbecker@med.uni-muenchen.de



Supplementary Figure 1: NSL1 antibody specificity control.
(A) Indirect immunofluorescence with the indicated antibodies in S2 cells treated with control and NSL1 RNAi. MSL1 (subunit of DCC complex) and lamin stainings are not compromised under NSL1 RNAi conditions. DNA was counterstained with DAPI. Scale bar (10 μ m) is indicated in white.
(B) Immunoblot analysis on S2 chromatin extracts (left) and Kc chromatin extracts (right) from NSL1 RNAi, and control RNAi (Glutathion-S-Transferase RNAi) using the NSL1 antibody. Lamin immunoblot was used as a loading control. The NSL1 protein shows aberrant mobility (11).



Supplementary Figure 2: NSL complex members bind similar across conditions.
(A-D) Scatter plots comparing NSL1 promoter binding in adult male and female flies (A), NSL1 vs. MBD-R2 in adult female flies (B, right), NSL1 binding with MBD-R2 binding in adult male flies (B, left), NSL1 binding and MBD-R2 binding (12) in S2 cells (C, left), MBD-R2 binding in S2 cells between the datasets of (12) and the modENCODE consortium (C, right), NSL1 binding in S2 cells with NSL1 binding in adult male cells (D). Rank-based Spearman correlation coefficients (ρ) are indicated.
(E) Genes were categorized in “housekeeping” and “differentially regulated” genes, respectively, based on the classification of Hurst and colleagues (22). Binding of NSL1 was displayed in these two groups and was found to be significantly higher in housekeeping genes than differentially regulated genes (Welch two sample t test, $p < 2.2 \times 10^{-16}$). Similar results were obtained using the S2 MBD-R2 ChIP-chip dataset from (12) and the MBD-R2 dataset generated by the modENCODE consortium (17). Note that the “housekeeping” definition after Hurst and colleagues is more stringent, which results in a number of genes classified as “housekeeping” by our definition but “differentially regulated” by Hurst and colleagues.

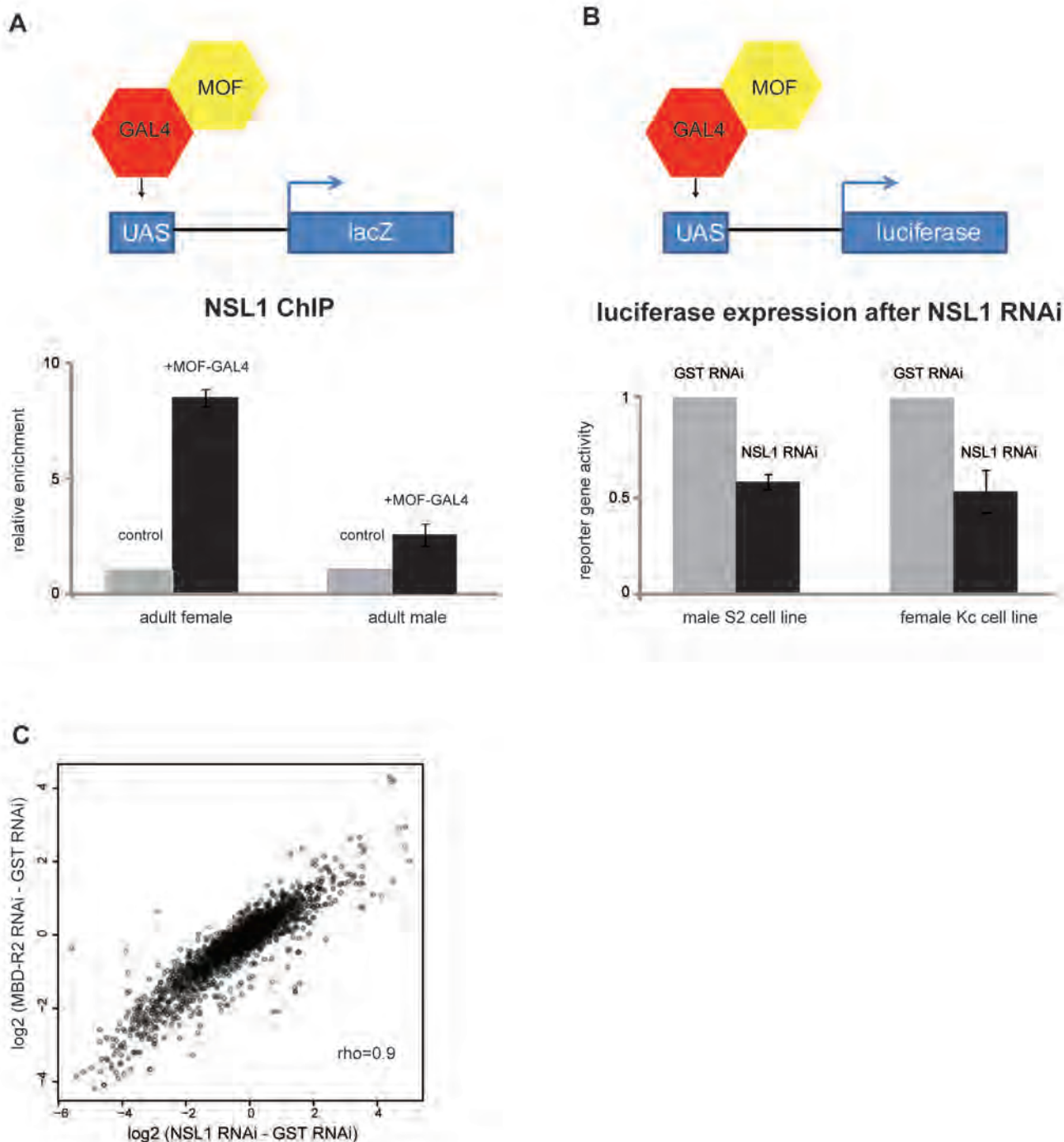
Supplementary Figure 3: Gene Ontology (GO) enriched biological processes of NSL complex bound genes

Term	PValue
GO:0051188~cofactor biosynthetic process	2,33E-12
GO:0031023~microtubule organizing center organization	7,03E-11
GO:0051186~cofactor metabolic process	1,40E-03
GO:0007017~microtubule-based process	1,70E-03
GO:0051297~centrosome organization	1,70E-03
GO:0009108~coenzyme biosynthetic process	3,00E-03
GO:0007098~centrosome cycle	4,50E-03
GO:0000226~microtubule cytoskeleton organization	4,90E-03
GO:0070271~protein complex biogenesis	5,80E-03
GO:0006461~protein complex assembly	5,80E-03
GO:0006732~coenzyme metabolic process	9,70E-03
GO:0016071~mRNA metabolic process	9,80E-03
GO:0043069~negative regulation of programmed cell death	9,90E-03
GO:0060548~negative regulation of cell death	9,90E-03
GO:0006397~mRNA processing	1,03E-02
GO:0022402~cell cycle process	1,17E-02
GO:0000278~mitotic cell cycle	1,17E-02
GO:0051656~establishment of organelle localization	1,50E-02
GO:0033043~regulation of organelle organization	1,55E-02
GO:0007010~cytoskeleton organization	1,82E-02
GO:0006352~transcription initiation	2,25E-02
GO:0043066~negative regulation of apoptosis	2,26E-02
GO:0022403~cell cycle phase	2,40E-02
GO:0007049~cell cycle	2,46E-02
GO:0051298~centrosome duplication	2,63E-02
GO:0007173~epidermal growth factor receptor signaling pathway	2,69E-02
GO:0000279~M phase	2,76E-02
GO:0051130~positive regulation of cellular component organization	2,99E-02
GO:0009314~response to radiation	3,30E-02
GO:0017038~protein import	3,41E-02
GO:0033365~protein localization in organelle	3,44E-02
GO:0006367~transcription initiation from RNA polymerase II promoter	4,07E-02
GO:0006605~protein targeting	4,08E-02
GO:0051640~organelle localization	4,21E-02
GO:0006733~oxidoreduction coenzyme metabolic process	4,92E-02
GO:0048477~oogenesis	5,56E-02
GO:0016044~membrane organization	5,59E-02
GO:0006366~transcription from RNA polymerase II promoter	5,62E-02
GO:0007447~imaginal disc pattern formation	5,73E-02
GO:0007052~mitotic spindle organization	5,77E-02
GO:0006468~protein amino acid phosphorylation	5,88E-02
GO:0031124~mRNA 3'-end processing	6,00E-02
GO:0007276~gamete generation	6,06E-02
GO:0019953~sexual reproduction	6,74E-02
GO:0007051~spindle organization	6,99E-02
GO:0051276~chromosome organization	7,19E-02
GO:0008156~negative regulation of DNA replication	7,27E-02
GO:0007292~female gamete generation	8,03E-02
GO:0016310~phosphorylation	8,32E-02
GO:0007349~cellularization	8,38E-02
GO:0000289~nuclear-transcribed mRNA poly(A) tail shortening	8,59E-02
GO:0000288~nuclear-transcribed mRNA catabolic process, deadenylation-d	8,59E-02
GO:0035222~wing disc pattern formation	8,77E-02
GO:0051493~regulation of cytoskeleton organization	8,77E-02
GO:0009628~response to abiotic stimulus	9,11E-02
GO:0032504~multicellular organism reproduction	9,29E-02
GO:0048609~reproductive process in a multicellular organism	9,29E-02
GO:0051301~cell division	9,52E-02
GO:0006281~DNA repair	9,65E-02
GO:0065003~macromolecular complex assembly	9,66E-02
GO:0006259~DNA metabolic process	9,70E-02
GO:0016072~rRNA metabolic process	9,70E-02
GO:0006612~protein targeting to membrane	9,70E-02
GO:0006364~rRNA processing	9,70E-02

Gene Ontology (GO) enrichment analysis was performed for biological processes using DAVID GO SLIM (57). P values were derived from a Fisher Exact test using the DAVID web-implementation. All NSL1 bound genes were compared to all active genes in S2 cells.

Term	PValue	Term	PValue
GO:0007606~sensory perception of chemical stimulus	1,57E-07	GO:0007451~dorsal/ventral lineage restriction, imaginal disc	0,01073
GO:0055114~oxidation reduction	4,48E-07	GO:0048598~embryonic morphogenesis	0,01169
GO:0007600~sensory perception	3,79E-06	GO:0044275~cellular carbohydrate catabolic process	0,01187
GO:0050890~cognition	6,59E-06	GO:0046164~alcohol catabolic process	0,01187
GO:0007155~cell adhesion	1,49E-05	GO:0006508~proteolysis	0,01269
GO:0022610~biological adhesion	1,49E-05	GO:0007476~imaginal disc-derived wing morphogenesis	0,01297
GO:0035218~leg disc development	1,79E-05	GO:0007494~midgut development	0,01349
GO:0006952~defense response	2,93E-05	GO:0009996~negative regulation of cell fate specification	0,01349
GO:0048569~post-embryonic organ development	5,89E-05	GO:0042659~regulation of cell fate specification	0,01349
GO:0035286~leg segmentation	1,03E-04	GO:0010453~regulation of cell fate commitment	0,01349
GO:0035285~appendage segmentation	1,03E-04	GO:0010454~negative regulation of cell fate commitment	0,01349
GO:0046942~carboxylic acid transport	1,03E-04	GO:0007626~locomotory behavior	0,01432
GO:0015849~organic acid transport	1,03E-04	GO:0007472~wing disc morphogenesis	0,01436
GO:0006865~amino acid transport	1,35E-04	GO:0043473~pigmentation	0,01489
GO:0015837~amine transport	1,35E-04	GO:0007164~establishment of tissue polarity	0,0158
GO:0007166~cell surface receptor linked signal transduction	1,41E-04	GO:0001736~establishment of planar polarity	0,0158
GO:0007552~metamorphosis	1,44E-04	GO:0007435~salivary gland morphogenesis	0,01581
GO:0035111~leg joint morphogenesis	1,92E-04	GO:0022612~gland morphogenesis	0,01581
GO:0007444~imaginal disc development	2,05E-04	GO:0001654~eye development	0,01737
GO:0045087~innate immune response	3,64E-04	GO:0019318~hexose metabolic process	0,01837
GO:0007186~G-protein coupled receptor protein signaling pathway	3,97E-04	GO:0009072~aromatic amino acid family metabolic process	0,01867
GO:0048707~instar larval or pupal morphogenesis	3,99E-04	GO:0035167~larval lymph gland hemopoiesis	0,01867
GO:0048563~post-embryonic organ morphogenesis	4,28E-04	GO:0006935~chemotaxis	0,01867
GO:0007560~imaginal disc morphogenesis	4,28E-04	GO:0007469~antennal development	0,01867
GO:0042742~defense response to bacterium	5,24E-04	GO:0007440~foregut morphogenesis	0,01867
GO:0048737~imaginal disc-derived appendage development	5,54E-04	GO:0035166~post-embryonic hemopoiesis	0,01867
GO:0048732~gland development	5,60E-04	GO:0007608~sensory perception of smell	0,01867
GO:0009886~post-embryonic morphogenesis	6,30E-04	GO:0001738~morphogenesis of a polarized epithelium	0,02213
GO:0048736~appendage development	6,39E-04	GO:0006959~humoral immune response	0,02258
GO:0006955~immune response	6,55E-04	GO:0008219~cell death	0,02319
GO:0035214~eye-antennal disc development	7,53E-04	GO:0016265~death	0,02319
GO:0035114~imaginal disc-derived appendage morphogenesis	8,79E-04	GO:0007411~axon guidance	0,02345
GO:0007447~imaginal disc pattern formation	9,21E-04	GO:0035289~posterior head segmentation	0,02396
GO:0048565~gut development	9,88E-04	GO:0042688~crystal cell differentiation	0,02396
GO:0035107~appendage morphogenesis	0,00101	GO:0008587~imaginal disc-derived wing margin morphogenesis	0,02585
GO:0016052~carbohydrate catabolic process	0,00139	GO:0043067~regulation of programmed cell death	0,02652
GO:0009617~response to bacterium	0,00175	GO:0010941~regulation of cell death	0,02652
GO:0009791~post-embryonic development	0,00192	GO:0007478~leg disc morphogenesis	0,02674
GO:0050830~defense response to Gram-positive bacterium	0,00215	GO:0019730~antimicrobial humoral response	0,02721
GO:0007354~zygotic determination of anterior/posterior axis, embryo	0,00215	GO:0042060~wound healing	0,02983
GO:0035108~limb morphogenesis	0,00221	GO:0012501~programmed cell death	0,03034
GO:0060173~limb development	0,00221	GO:0009310~amine catabolic process	0,03038
GO:0035110~leg morphogenesis	0,00221	GO:0006007~glucose catabolic process	0,03224
GO:0007157~heterophilic cell adhesion	0,00238	GO:0019320~hexose catabolic process	0,03224
GO:0007449~proximal/distal pattern formation, imaginal disc	0,00238	GO:0048066~pigmentation during development	0,03293
GO:0035220~wing disc development	0,00258	GO:0016054~organic acid catabolic process	0,03293
GO:0007431~salivary gland development	0,00261	GO:0046395~carboxylic acid catabolic process	0,03293
GO:0035272~exocrine system development	0,00261	GO:0006576~biogenic amine metabolic process	0,03535
GO:0035120~post-embryonic appendage morphogenesis	0,00276	GO:0050829~defense response to Gram-negative bacterium	0,03535
GO:0042067~establishment of ommatidial polarity	0,00283	GO:0006334~nucleosome assembly	0,03535
GO:0010623~developmental programmed cell death	0,00378	GO:0035222~wing disc pattern formation	0,03789
GO:0007450~dorsal/ventral pattern formation, imaginal disc	0,00433	GO:0030030~cell projection organization	0,03796
GO:0007610~behavior	0,00438	GO:0000902~cell morphogenesis	0,03819
GO:0035215~genital disc development	0,00444	GO:0005976~polysaccharide metabolic process	0,03902
GO:0007219~Notch signaling pathway	0,00492	GO:0046365~monosaccharide catabolic process	0,03963
GO:0007366~periodic partitioning by pair rule gene	0,00509	GO:0042386~hemocyte differentiation	0,04036
GO:0001708~cell fate specification	0,00608	GO:0035162~embryonic hemopoiesis	0,04036
GO:0019731~antibacterial humoral response	0,00617	GO:0000272~polysaccharide catabolic process	0,04036
GO:0050877~neurological system process	0,00683	GO:0048568~embryonic organ development	0,04036
GO:0048190~wing disc dorsal/ventral pattern formation	0,00808	GO:0048749~compound eye development	0,04235
GO:0007362~terminal region determination	0,00834	GO:0046667~compound eye retinal cell programmed cell death	0,04283
GO:0050909~sensory perception of taste	0,00834	GO:0035161~imaginal disc lineage restriction	0,04283
GO:0009954~proximal/distal pattern formation	0,00834	GO:0009253~peptidoglycan catabolic process	0,04283
GO:0006006~glucose metabolic process	0,00866	GO:0006027~glycosaminoglycan catabolic process	0,04283
GO:0006928~cell motion	0,00924	GO:0000270~peptidoglycan metabolic process	0,04283
GO:0016337~cell-cell adhesion	0,00933	GO:0006206~pyrimidine base metabolic process	0,04283
GO:0002165~instar larval or pupal development	0,00964	GO:0035287~head segmentation	0,04283
GO:0009065~glutamine family amino acid catabolic process	0,01053	GO:0035223~leg disc pattern formation	0,04283
GO:0006584~catecholamine metabolic process	0,01053	GO:0007479~leg disc proximal/distal pattern formation	0,04283
GO:0009636~response to toxin	0,01053	GO:0014070~response to organic cyclic substance	0,04416
GO:0040034~regulation of development, heterochronic	0,01053	GO:0007419~ventral cord development	0,04416
GO:0018958~phenol metabolic process	0,01053	GO:0043279~response to alkaloid	0,04416
GO:0034311~diol metabolic process	0,01053	GO:0006096~glycolysis	0,04557
GO:0009712~catechol metabolic process	0,01053	GO:0045596~negative regulation of cell differentiation	0,0459
		GO:0042981~regulation of apoptosis	0,04595
		GO:0043068~positive regulation of programmed cell death	0,04734
		GO:0030031~cell projection assembly	0,04734
		GO:0010942~positive regulation of cell death	0,04734

Gene Ontology (GO) enrichment analysis was performed using active but NSL1 unbound genes against the background of all active genes. DAVID GO SLIM (57) terms were used for biological processes. P values were derived from a Fisher Exact test using the DAVID web-implementation.



Supplementary Figure 5: NSL1 stimulates transcription on a reporter gene and genome-wide.

(A) NSL1 binding correlates with stronger activation in female flies as compared to male flies after MOF recruitment to a reporter gene. We previously showed that recruitment of the MOF-GAL4 fusion protein to an UAS-element upstream of the lacZ reporter gene leads to stronger activation in adult female flies as compared to adult male flies (12). MBD-R2 recruitment as measured by ChIP at the lacZ gene follows this trend (12). Here, we tested NSL1 binding at the reporter gene and find a similar recruitment of NSL1 which is significantly ($p < 0.05$, 2-sided unpaired t-test) increased in adult female flies when compared to adult male flies. We conclude that the higher activation potential of MOF in female flies correlates with an increased NSL complex recruitment. Error bars represent SEM of at least three independent biological replicates.

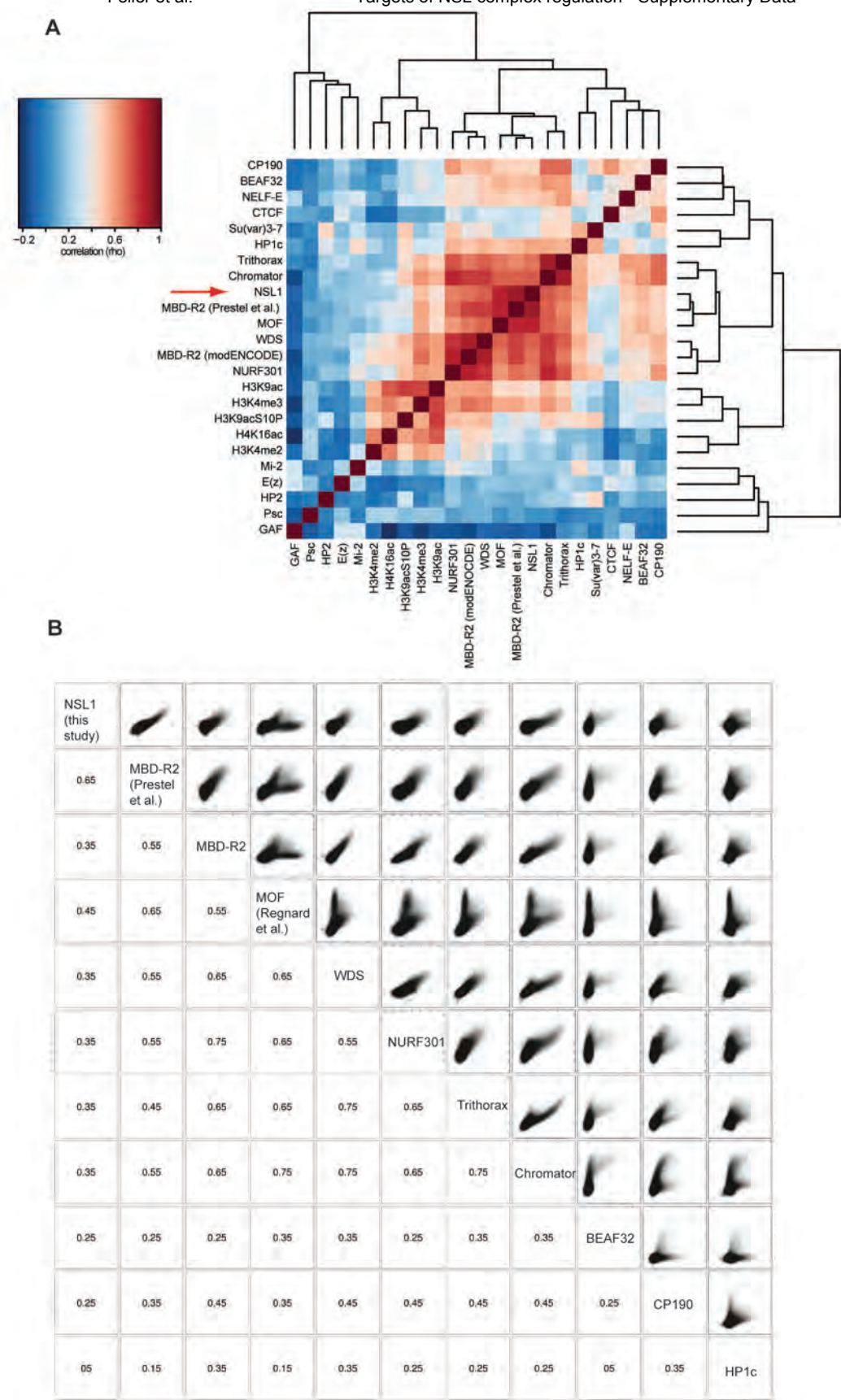
(B) RNAi mediated ablation of NSL1 leads to a diminished MOF-dependent activation of a luciferase reporter gene in male S2 cells and female Kc cells. Transiently co-transfection of MOF-GAL4 fusion gene, UAS-luciferase reporter and renilla reporter (used for normalization) lead to an activation of the luciferase reporter. NSL1 RNAi in male S2 and female Kc cells significantly diminished ($p < 0.05$, 2-sided unpaired t-test) the luciferase reporter expression when compared to a control RNAi treatment. Error bars represent SEM of at least three independent biological replicates.

(C) Transcriptome changes are similar in response to NSL1 RNAi and MBD-R2 RNAi. Scatter plot of transcriptome changes after NSL1 (this study) and MBD-R2 RNAi (12), respectively, relative to control RNAi. Rank-based Spearman correlation coefficient (ρ) is indicated in the corner.

Supplementary Figure 6: Gene Ontology enrichment analysis of NSL1 RNAi downregulated genes

Term	Pvalue
GO:0006259~DNA metabolic process	1,30E-05
GO:0006281~DNA repair	2,70E-04
GO:0034660~ncRNA metabolic process	9,00E-04
GO:0006412~translation	7,23E-02
GO:0033554~cellular response to stress	0,001
GO:0006399~tRNA metabolic process	0,001
GO:0006974~response to DNA damage stimulus	0,001
GO:0008033~tRNA processing	0,002
GO:0034470~ncRNA processing	0,005
GO:0006284~base-excision repair	0,006
GO:0006366~transcription from RNA polymerase II promoter	0,011
GO:0006367~transcription initiation from RNA polymerase II promoter	0,012
GO:0006351~transcription, DNA-dependent	0,014
GO:0006352~transcription initiation	0,015
GO:0032774~RNA biosynthetic process	0,016
GO:0006396~RNA processing	0,017
GO:0009451~RNA modification	0,023
GO:0006260~DNA replication	0,032

Downregulated genes (fold change < -2) were compared to all active genes (Affymetrix log2-scale > 4). Gene Ontology enrichment analysis was performed using DAVID (57), testing for biological processes using the GO SLIM terms. P values were derived from a Fisher Exact test using the DAVID web-implementation. A similar result was obtained using the R/Bioconductor package GStats (data not shown).

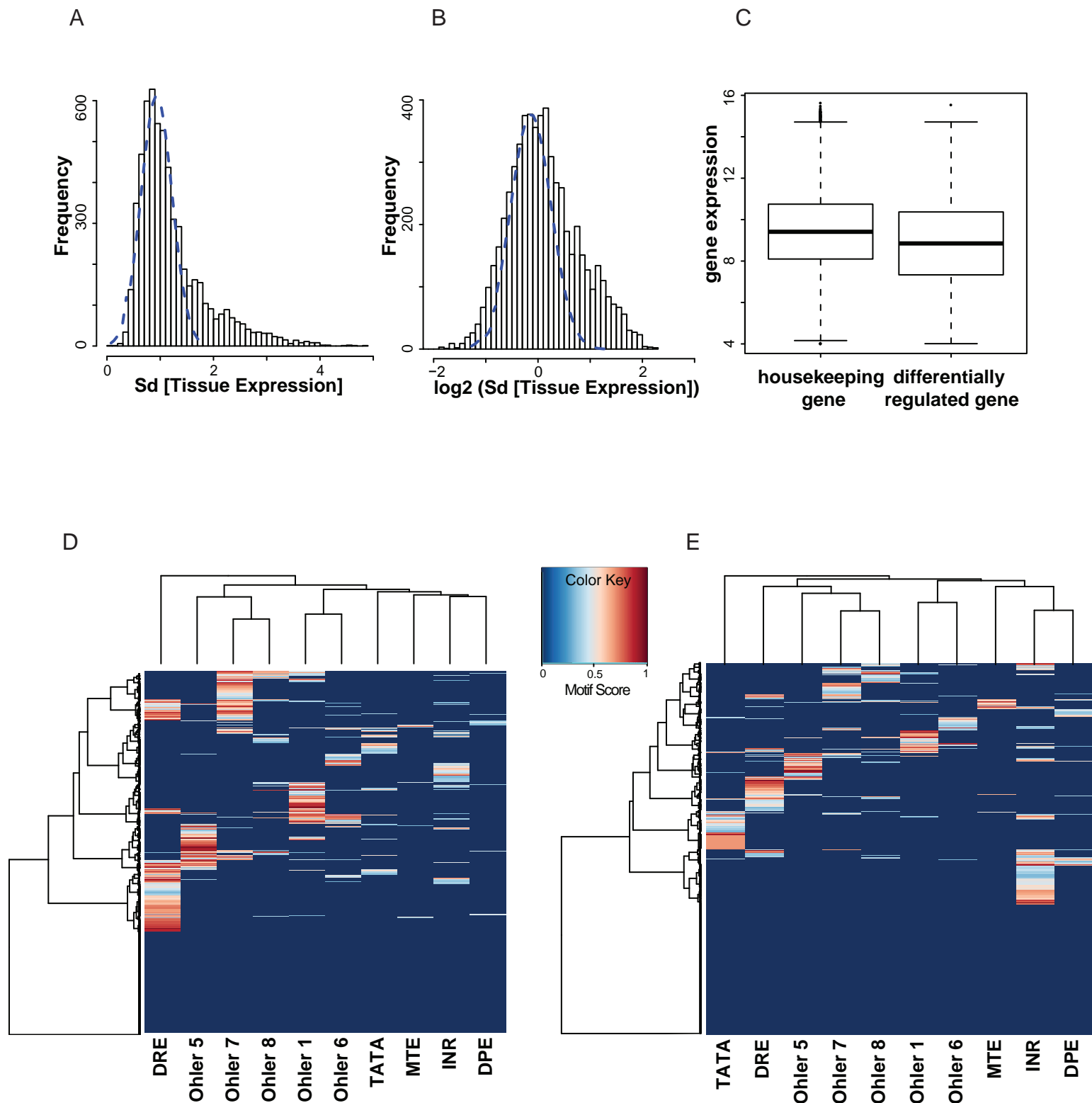


Supplementary Figure 7: Co-localization of NSL complex members with Trithorax and NURF301.
(A) Heatmap was generated as in Figure 3A but included in addition to the profiles shown in Figure 3A the ChIP-chip datasets of MBD-R2 (12), MBD-R2 (modENCODE) and MOF (modENCODE).
(B) Probe-wise scatter plot indicated co-localization of the NSL complex with Trithorax and NURF301. ChIP-chip datasets were used from the modENCODE consortium database (17) unless otherwise indicated. Probe-wise enrichments were calculated, smoothed and genomic positions were compared in 100 base pair intervals across samples. Note that all probes are shown including the inactive unbound probes and the X chromosomal active probes where NSL1 and MBD-R2 binding deviate from MOF binding towards the transcribed regions of genes (compare to (12)).
Spearman correlation coefficient is indicated.

Supplementary Figure 8: Summary of factor and sequence occupancies across different conditions.

	% Binding to all			% Binding to all		Enrichment	Enrichment
	Genes	Active	Inactive	active	active	Active/	Housekeeping/
		Genes	Genes	Housekeeping	DRG	Inactive	/DRG
BEAF32	17,9	33,5	1,1	39,9	17	31,7	2,3
Ohler 1	7,6	12,7	2,1	15	6,7	5,9	2,2
HP2	2	3,2	0,7	3,8	1,7	4,8	2,2
DRE	14,4	21,2	7,2	24,6	12,3	2,9	2
Ohler 5	8,6	12,1	4,7	14	7,2	2,6	2
Chromator	44,7	80,7	6	92,7	49,4	13,4	1,9
Ohler 7	10,4	15,3	5,1	17,6	9,5	3	1,9
MBD-R2	38,9	71,7	3,7	82,1	44,7	19,4	1,8
WDS	39,6	73,5	3,1	83,9	46,7	23,5	1,8
Trithorax	40,6	74,4	4,4	84,7	47,5	16,7	1,8
CP190	29	47,5	9,2	54	30,5	5,2	1,8
NURF301	42,2	77,5	4,2	87,8	51,1	18,4	1,7
Ohler 6	6	7,6	4,2	8,5	5,1	1,8	1,7
HP1c	15,5	29,1	1	32,6	20	29,1	1,6
H3K4me3	47,5	87,1	5,1	94,9	67	17,2	1,4
Su(var)3-7	1,8	2,6	1,1	2,8	2	2,4	1,4
H3K9ac	46,9	84,5	6,6	92	65,3	12,8	1,4
Nelf-E	28,8	54,9	0,9	59,3	43,3	59,6	1,4
H4K16ac	50,0	80,0	17,5	78,9	58,6	9,7	1,3
Nelf-B	30,7	58,6	0,9	62,7	47,7	62,3	1,3
CTCF	4	4,4	3,5	4,5	4,2	1,2	1,1
Ohler 8	5,8	6,8	4,7	6,9	6,6	1,4	1
TATA	9,7	8,2	11,4	7,9	8,9	0,7	0,9
E(z)	1,3	0,8	1,9	0,7	1,1	0,4	0,6
INR	18,4	15,6	21,5	12,7	23,2	0,7	0,5
H3.3	20,8	32,3	8,4	26,2	48,3	3,8	0,5
MTE	2	1,9	2,2	1,5	3	0,9	0,5
Psc	1,6	1,1	2	0,8	1,9	0,6	0,4
RING	2,9	3,3	2,4	2,4	5,8	1,4	0,4
DPE	3,6	3,4	3,8	2,5	6	0,9	0,4
Mi-2	2,6	4,7	0,3	2,9	9,2	15,8	0,3
GAF	8,3	14,8	1,4	8,6	30,7	10,6	0,3
H3K27me3	35,8	4,1	69,8	1,9	9,8	0,1	0,2
Pc	2	0,5	3,6	0,2	1,2	0,1	0,1

Factor and sequence occupancy is given relative to all genes in the *D. melanogaster* genome (FlyBase dmel5.32 annotation) (column 2), all active genes (column 3, gene activity measured by modENCODE polymerase II dataset), all inactive genes (column 4), all active housekeeping genes (column 5) and all active “differentially regulated” genes (DRG, column 6). Column 7 indicates the ratio of active genes over inactive genes, and column 8 indicates the ratio of active housekeeping genes over active “differentially regulated” genes. The list was sorted according to the enrichment towards active housekeeping genes (column 8) and overrepresented and underrepresented factors are labeled in red and green, respectively. Only modENCODE ChIP-chip profiles were used.



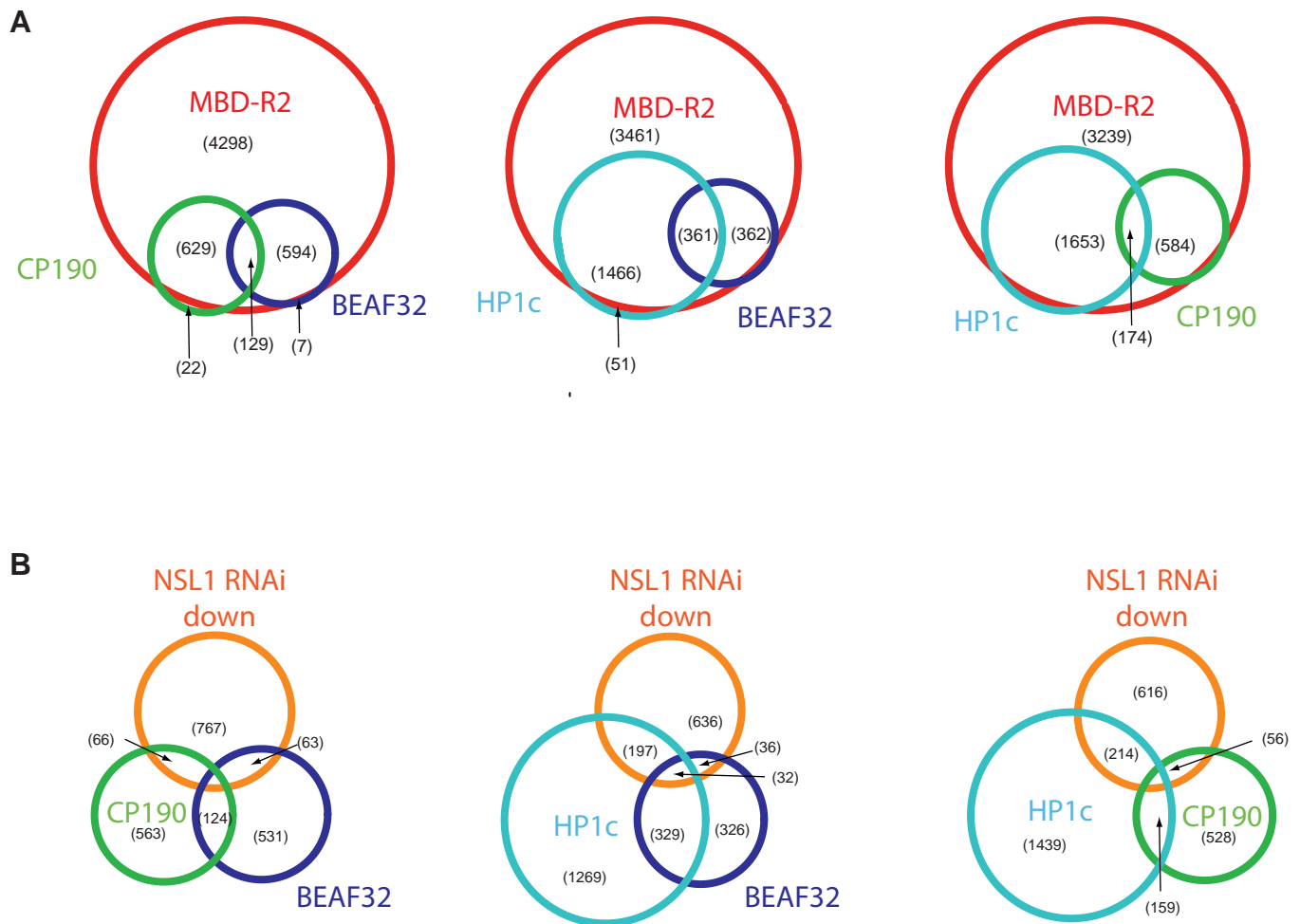
Supplementary Figure 9: Characterization of core promoter motifs at active genes.

(A) Histogram depicting the distribution of standard deviations of active genes across 40 *Drosophila* tissues (21). Normal distribution is overlaid with a blue dashed line. (B) Same as (A) but \log_2 transformed standard deviations. (C) Comparison of gene expression values in housekeeping genes (left box) and differentially regulated genes (right box). Gene expression values were taken from the \log_2 transformed Affymetrix control RNAi transcriptome dataset. (D) Core promoter motif distribution of active housekeeping genes. Heatmap presentation of motif score (columns) along active housekeeping genes (rows). The dendrogram to the right was generated by hierarchical clustering of gene promoters according to their motif score using the 'complete' method implemented in the R package 'hclust'. The dendrogram above the heatmap was generated by hierarchical clustering ('ward method') of the Spearman correlated motif scores. Note that 1269 of the 4523 active housekeeping genes do not show any significant signature of at least one of the investigated core promoter motifs. See Material and Methods for detailed information on the definition of the 'motif score'. (E) Same as (D) but for active differentially regulated genes. Note that 480 of the 1367 active differentially regulated genes do not show any significant signature of at least one of the investigated core promoter motifs.

Supplementary Figure 10: NSL1 peaks are centered over core promoter motifs characteristic for housekeeping genes

motif	total number sequences enriched for motif	sequence fraction enriched for motif
DRE	1106	19.9
Ohler 1	703	12.7
Ohler 5	556	10.0
Ohler 7	510	9.2
Ohler 6	251	4.5
Ohler 8	109	2.0
INR	84	1.5
TATA	38	0.7
MTE	33	0.6
DPE	12	0.2

De-novo sequence search (Hartmann and Soeding, in preparation) of the sequence surrounding +/- 50 base pairs centered on the NSL1 ChIP-Seq (11) peak summit. Column 2 indicates the number of identified sequences containing the respective motif. Column 3 displays the fractions of each identified motif relative to all identified motifs.



Supplementary Figure 11: HP1c and insulator binding proteins co-localize with the NSL complex at non-functional sites.

(A) Venn diagram representation of gene sets bound by the NSL complex (using the MBD-R2 ChIP-chip profile generated by the modENCODE consortium) and heterochromatin protein 1c (HP1c), CP190 and BEAF32. (B) Venn diagram representation of gene sets downregulated after NSL1 RNAi and bound by the heterochromatin protein 1c (HP1c), CP190 and BEAF32.

Numbers in brackets represent the number of genes in the respective group (for clarity numbers with little information content are not shown). We used the modENCODE MBD-R2 data set as a proxy for the NSL complex to minimize technical variation in the comparison with the BEAF32/CP190/HP1c datasets. Similar results were obtained using the NSL1 dataset and the MBD-R2 dataset (12), respectively.

3.4 Global and specific responses of the histone acetylome to systematic perturbation

Christian Feller, Ignasi Forné, Axel Imhof and Peter B. Becker

3.4.1 Summary, significance and own contribution

Summary and significance

Post-translational protein modifications (PTMs) are at the core of most regulatory systems. Histone proteins acquire impressive patterns of PTMs that govern all nuclear processes and often indicate or even cause malfunction that manifests in complex diseases including cancer and neurodegenerative diseases.

Although research in epigenetics and chromatin over the past two decades documented many, mostly *qualitative*, relationships between *individual* histone PTMs and the enzymes that write and erase these marks, a *quantitative* and *comprehensive* description of all enzyme-substrate relationships for an entire class of PTMs is missing but urgently needed. Such a global description of PTMs and their modifiers as a *system* will greatly facilitate our ability to develop quantitative models on chromatin pathways that are not only beneficial to refine our understanding of nuclear processes and their malfunction but also to develop new epigenetic therapies.

In this study, we generated a quantitative and comprehensive catalogue of histone acetylation and methylation motifs that describe the changes in response to ablation of each single known or suspected lysine acetyltransferase (KAT) and deacetylase (KDAC) in *Drosophila melanogaster*.

The comprehensiveness of the dataset not only allows us to define *class-specific* features that are shared among members of enzyme families but also to describe the histone modification system as a highly *interconnected, responsive network* that compensates for the loss of single components.

Generating this catalogue first required optimization of liquid chromatography mass spectrometry (LC-MS) workflows. Our improved LC-MS protocol not only allowed precise and accurate quantification of *combinatorial* histone acetylation and methylation motifs, but also delivered a general applicable tool to identify and quantify combinatorial PTM motifs on other proteins. We describe an inherent bias common to most previous histone PTM studies applying LC-MS and provide a resource to correct for this bias.

Our classification of acetylation motifs according to their absolute abundances has direct implications on the design and interpretation of genomic experiments.

The most surprising finding was that ablation of almost every KAT triggered a systemic response to effectively maintain global acetylation levels. Its widespread occurrence and observation in different species, including human cells, suggests this to be a fundamental, conserved principle. Considering the *global* but distinct changes to the epigenome when inhibiting *individual* KAT activities may be utilised to improve current and develop more effective epigenetic therapies.

Applying the methodology to another, already well-documented ‘compensatory system’ revealed that chromosome-wide transcriptional fine-tuning during *Drosophila* dosage compensation is not restricted to the change of a single histone PTM (H4.K16ac), but accompanied by an global yet specific re-distribution of acetylation and methylation marks. This finding provides further evidence that a full understanding of the function of a *single* histone PTM requires to analyse it in the context of the entire *system*.

In summary, our data not only provides a general resource for enzyme-substrate relationships to KATs and KDACs but it also defines a starting point to explore systematically the interconnected network properties of chromatin modification pathways.

Own contribution

Prof. Peter Becker and I conceived the project and wrote the manuscript. I co-developed the LC-MS workflow with Dr. Ignasi Forne. I performed all experiments, analysed the data and prepared all figures. Most peptide samples were processed by Dr. Ignasi Forne, who also performed the MS measurements. Prof. Axel Imhof supervised the LC-MS experiments and contributed important reagents.

3.4.2 Submitted manuscript

Global and specific responses of the histone acetylome to systematic perturbation

SUMMARY

Regulation of histone acetylation is fundamental to the utilization of eukaryotic genomes in chromatin. Aberrant acetylation contributes to disease and can be clinically combated by inhibiting the responsible enzymes.

Our knowledge of complex histone acetylation patterns is patchy because current methodologies do not permit a systematic assessment of rare and combinatorial signatures. We devised a generally applicable, sensitive mass spectrometry-based strategy to efficiently and quantitatively monitor combinatorial modifications. This was applied to generate a comprehensive inventory of acetylation motifs on histones H3 and H4 in *Drosophila* cells.

Systematic depletion of every known or suspected acetyltransferase and deacetylase revealed specific alterations of histone acetylation signatures, established enzyme-substrate relationships and unveiled an extensive crosstalk between neighboring modifications. Unexpectedly, overall histone acetylation levels remained remarkably constant upon depletion of individual acetyltransferases. Conceivably, the acetylation level is adjusted to maintain the global charge neutralisation of chromatin and the stability of nuclei.

Highlights

- System's view of histone modification motifs and its response to enzyme depletion
- Extensive crosstalk between lysine acetylation and methylation
- Global acetylation levels are maintained in response to acetyltransferase ablation
- Global chromatin features associated with fly dosage compensation

INTRODUCTION

The problem of packaging and organising complex genomes in the nuclei of cells was elegantly solved by the evolution of chromatin. The wrapping of DNA around octamers of histone proteins to form an ever-repeating succession of nucleosomes is universal among all eukaryotes. The increasing complexity of organisms due to the differentiation of cell types, necessitates a structural and functional diversification of chromatin. This is mainly achieved by post-translational, chemical modifications of histones.

Currently, we know of more than 10 different types of chemical modifications (such as methylation, acetylation, phosphorylation) that alter the properties of amino acids in histones. Collectively, these modifications locally define chromatin structure and are fundamental to establishing the gene expression programmes that characterise healthy and diseased cells. The enzymes that reversibly modify histones are increasingly recognised as targets for therapeutic intervention in cancer tissue and neurodegenerative diseases (Dawson and Kouzarides, 2012; Graff and Tsai, 2013).

Acetylations of histones H3 and H4 are frequent and among the best characterised post-translational modifications (Sternier and Berger, 2000; Kouzarides, 2007; Lee and Workman, 2007). The N-terminal domains of these two histones (the amino-terminal ‘tail’ domains) alone bear a dozen lysine residues subject to acetylation. Additional acetylation has also been reported at internal residues (Tan et al., 2011).

The acetylation of lysine has structural consequences: charge neutralization weakens DNA-histone and nucleosome-nucleosome interactions and is often correlated to unfolding of nucleosome fibres and gene activation (Kouzarides, 2007). Acetylated lysines may contribute to marking individual histones as they are recognised by bromodomains of effector proteins. Acetylation of lysines prevents their methylation or ubiquitylation and may therefore have secondary effects. Histone acetylation patterns in chromatin result from the interplay between an impressive number of dedicated lysine acetyltransferases (KATs) and antagonizing deacetylases (KDACs).

In evaluating the potential of an individual acetylation mark, it is useful to know its abundance. For example, a highly abundant acetylation is likely to have a broad, structural function. By contrast, a rare acetylation might rather serve as a signal. Currently, our knowledge on the cellular copy number for most histone modifications is sparse.

Different post-translational modifications (PTMs) may reside on the same histone molecule. Many cases have been reported where PTM combinations (PTM motifs) rather than individual marks bear functional meaning (Fischle et al., 2003). Top-down mass spectrometry suggests the existence of over 200 different PTM motifs on histones – many of which involve site-specific acetylation – yet their exact cellular abundance remains elusive (Pesavento et al., 2008; Young et al., 2009; Jung et al., 2013).

Acetylated lysines may be bound by dedicated bromodomains in effector proteins. Although the affinity towards a single acetyl-lysine is low, multiple adjacent acetyl-lysines not only boost the affinity but may also tune the targeting selectivity (Ruthenburg et al., 2007; Filippakopoulos et al., 2012). It is currently not known whether combinatorial acetylation motifs are generated by a single KAT or whether they require the action of multiple enzymes, with the exception of the cytoplasmic deacetylation of H4 at K5 and 12, which is catalysed by HAT1 (Parthun, 2012).

Moreover, PTMs neighbouring acetylated lysines can strongly alter the affinity of bromodomains – and likewise the enzymatic activity of KATs and KDACs may be modulated by close-by PTMs (Fischle et al., 2003; Filippakopoulos et al., 2012).

Our ability to detect low abundant histone acetylation, or acetylation as part of a PTM motif are currently rather limited. Traditionally, histone PTMs are detected with antibodies raised against appropriately modified peptides. However, recent studies alarm that antibodies raised against individually acetylated lysines (e.g. H4.K12ac) primarily recognise multiply-acetylated peptides (e.g. tetra-acetylated H4) and may display substantial off-target reactivity (Suka et al., 2001; Fuchs et al., 2011; Rothbart et al., 2012). Moreover, antibody detection systems have a limited linear dynamic range, precluding a quantitative analysis. Finally, most antibodies are unable to differentially display acetylation in combinatorial motifs.

Peptide modifications are commonly detected by mass spectrometry (MS), however, most current MS protocols are hampered by low sensitivity and dynamic range and high-level requirements for operating personnel and instrumentation. We optimised existing procedures and developed new protocols to establish a robust and generally applicable workflow for accurate and precise quantification of modification motifs. We applied this methodology to generate a comprehensive inventory of motifs involving acetylation of H3 and H4 in *Drosophila* cells. This database will become a useful tool for further studies.

In order to learn how acetylation motives are generated we systematically assessed the effects of individual acute depletion of all putative KAT and KDAC enzymes in *Drosophila*. This effort identified new activities and specificities and re-defined enzyme-substrate relationships that had been controversial or simply unknown. To our surprise, we also uncovered a global response of the histone acetylome to interference with certain acetylation pathways, which suggests that the total concentration of histone acetylation in nuclei is subject to homeostatic control.

Our classification of histone acetylation motifs according to their abundance, enzyme-dependence and their contribution to the systemic response to acute perturbation of specific acetylation reactions leads to testable hypotheses about their contributions to global chromosome architecture and local signalling events.

RESULTS

Optimizing liquid chromatography mass spectrometry workflows for accurate and precise quantification of single and combinatorial PTMs

Identifying and quantifying rare, combinatorial modification motifs currently poses a major analytical challenge. We optimised liquid chromatography mass spectrometry (LC-MS) workflows for sensitive, precise and accurate quantification of PTM motifs containing lysine acetylation and methylation.

We rapidly isolated histones from cells by acid extraction and chemically acetylated all unmodified and mono-methylated lysines with deuterated acetic anhydride (D3AA method (Smith et al., 2003)). The deuterated (d3) mass tag adds three Daltons to distinguish endogenous from chemical acetylation. This also ensures that digestion by trypsin only occurs after arginine, yielding peptides of an intermediate size necessary for bottom-up analysis of PTM motifs (Figure 1A). Unlike the more common propionylation (Garcia et al., 2007; Schotta et al., 2008; Zheng et al., 2013), chemical acetylation yields peptides with similar physicochemical properties to those bearing endogenous acetylation. This ensures comparability during the entire LC-MS workflow, including the MS quantification.

We developed a generally applicable targeted MS workflow to increase the dynamic range of MSⁿ measurements (see Suppl. Note 1). This was necessary because many combinatorial motifs can only be resolved by consecutive MS1 and MS2 scans, and four H4 di-acetyl motifs require MS1-MS2-MS3 (see below). The targeted approach increased the precision and accuracy to quantify minor differences in rare and combinatorial histone PTMs (Figures 1B, C, S1, S2, S3 and Suppl. Note 1).

Motifs containing methylated lysines in addition to acetylation displayed characteristic shifts in retention time (Figures 1A, S1). A systematic analysis of positional isomers (peptides with identical mass but PTMs at different positions) revealed some rules. A di- or tri-methyl group close to the N-terminus of a peptide results in a lower hydrophobicity and thus earlier retention time as compared to when the methylation is in the middle. For example, an H3 peptide containing tri-methylated lysine 27 (K27me3) elutes a few minutes before the same peptide bearing K36me3 (Figure S1B). Furthermore, tri-methylated peptides elute shortly before di-methylated ones, followed by unmethylated and, lastly, mono-methylated forms. This was true for all histone peptides analysed and allows a precise quantification of peptides carrying methylation-only or mixed acetylation-methylation motifs solely based on their MS1 values (Figures 1A, S1).

Positional isomers carrying one or multiple acetylations are not separated chromatographically using the D3AA method. Their individual intensities can be quantified using signals generated by a high occurrence of targeted MS2 scans (Figures 1A, S2; Tables S1, S2). Because four of the di-acetylated histone H4 combinations, including H4.K5acK12ac, cannot be identified and quantified with MS1 and MS2 only, we optimised procedures to quantitatively read out prognostic MS3 ions and thereby

measured those isoforms by successive MS1-MS2-MS3 scans (Figures S2A, S2B, Suppl. Note 1.2). To our knowledge, this is the first direct MS-based quantification of all six histone H4 di-acetylated isoforms.

To validate and benchmark the LC-MS workflow, we used synthetic peptides carrying different acetylation motifs at different ratios (Figure S2B, Table S3), which demonstrated a very high accuracy. To assess the level of precision, we performed three whole-workflow replicates starting with 2 million KC cells each. The quantification of 45 histone motifs resulted in a median coefficient of variation (CV) of 5.2% (Figure 1B, Table S4). Next, we measured the level of variation for five independent biological replicates involving control RNAi treatments of between 2 and 8 million cells. Through optimising MS2 and MS3 recordings, the CVs for motifs, which require successive MSⁿ measurements were similar (median CV of 10.4) to those for motifs requiring MS1 only (median CV of 8.9). However, the CVs of the individual histone PTMs varied between 0.8% and 35.7% (Figure 1B, Table S4). Similar results were obtained using 75,000 to 300,000 mammalian cells (data not shown). The differences in CV for histone PTMs reflect their *in vivo* abundances, which differ in several orders of magnitude (Figures 1D, E) and the technical imprecision specific to any MS identifier (Table S4). For example, quantitation of the highly abundant H3.K27me2 motif requires only MS1-based quantification with highly precise MS1 quantifier (CV = 5.3%), whereas the low-abundant H4.K8acK12ac motif requires integration from MS1, MS2 and MS3 peak areas (CV = 35.7%). Overall, our current LC-MS workflow is able to analyse low amounts of sample and the sensitivity, precision and accuracy are improved over state-of-art workflows for quantifying PTM motifs.

Correction of LC-MS signal bias

Modifications change the physiochemical properties of peptides, which may influence several steps during the LC-MS workflow and thereby impair accurate quantification (Marx et al., 2013). To assess this phenomenon, we determined the LC-MS response factor by measuring a library of synthetic peptides containing acetylation and methylation motifs to H3 and normalising the proteotypic signals relative to the Q tag abundance (Figure 1C, Suppl. Note 2). Equimolar solutions of peptides with different methylation signatures display very drastic differences in the LC-MS response factor, corroborating a recent study by Garcia and colleagues (Lin et al., 2014). The effect is most dramatic for peptides containing H3.K9me3, whose Q tag normalised LC-MS response factor is 34 times lower than that of the K9me1 peptide. Therefore, the interpretation of the cellular abundance of K9me3 is underestimated by an order of magnitude if not corrected for.

On the other hand, H3 peptides bearing K9ac or K14ac have almost identical LC-MS response factors, confirming that the D3AA method achieves accurate quantification of acetylation motifs.

In summary, experimentally determined LC-MS response factors allowed introducing a correction factor specific to histone motifs. Table S3 summarises these factors for a selected set of peptides. Our catalogue of correction factors will facilitate the interpretation of published and future PTM datasets for relative and absolute abundances of peptides and their PTMs.

The histone acetylation motifs fall into three abundance classes

We measured the acetylation and methylation levels of all lysines on histone H3 and H4 in *Drosophila melanogaster* KC cells and quantified 45 of over 55 motifs detected (Figure 1D, Table S4, Suppl. Note 1.3). This included acetylation of K5, K8, K12 and K16 as well as methylation of K20 on histone H4 and seven acetylation (K9, K14, K18, K23, K27, K36, K37) as well as five methylation sites (K4, K9, K27, K36, K79) on H3. We did not detect acetylation or methylation on the remaining six lysines of histone H4 and four lysines on histone H3. Although other lysine acetylation sites had been reported before (H4.K91ac, H3.K4ac, H3.K56ac, H3.K122ac), their abundance in *Drosophila* cells must be below our detection limit of approximately 0.003% or 182 histone molecules per cell, at least on average in an asynchronous population (see below and Suppl. Note 1.3).

Knowledge of the cellular abundances of histone motifs facilitates the interpretation of chromatin maps and may contribute to developing quantitative models for the function of histone modifications. For example, H3.K9me3 constitutes a hallmark of repressive chromatin that covers significant regions of the fly and mammalian genomes (Kharchenko et al., 2011; Consortium et al., 2012). In support of this widespread notion, we measured K9me3 on 39% of H3 molecules, but only if we corrected for the differential LC-MS response (without correction: 4.2% H3.K9me3, Table S4).

The abundances among individual acetylation motifs differ by four orders of magnitude (Figure 1D). Plotting the abundance distribution of the 29 histone acetylation motifs revealed three abundance classes. H3.K23ac is by far the most abundant acetylation (47% of H3, corresponding to 2.8 million molecules/cell). It is present, on average, on every second nucleosome or even on all nucleosomes if the arrangement was heterotypic. Eleven motifs were of intermediate abundance (1–12% of histones, class median: 3.7%). Most combinatorial acetylation motifs, however, are present on less than 1% of histones (17 motifs, class median: 0.3%). For example, the H4.K5acK8acK16ac motif is present in only about 7600 estimated copies per cell (for reference, a KC cell has about 6000 active gene promoters). It is tempting to speculate that highly abundant histone acetylation motifs contribute to global features of chromatin structure, whereas rare and combinatorial histone acetylation motifs constitute marks for bromodomain-mediated signalling.

The contributions of all known or suspected KAT and KDAC enzymes to histone acetylation motifs

The inventory of all histone acetylation motifs serves as a point of entry to study the histone acetylome as a structural and regulatory system that is installed by KATs and KDACs according to cell-specific programmes. The properties of this system depend on the substrate specificity or redundancy of the enzymes. Which KATs and KDACs contribute to a specific histone acetylation motif? What are the global effects on the histone acetylation landscape if a single enzyme is depleted? Addressing these questions we systematically evaluated the contribution of every KAT and KDAC to the inventory.

Drosophila melanogaster is particularly suited for a single gene perturbation strategy, since most KATs, KDACs and KAT/KDAC-associated proteins are well conserved between the fly and human cells, yet, the fly genome does not contain extensive KAT and KDAC paralogues, which are common for most mammalian genomes.

We created and curated a list of putative acetyltransferases and deacetylases in *D. melanogaster*, including novel candidate genes coding for proteins with acetyltransferase-related domains, using public databases, primary research articles and cross-homology search using BLAST (Table S6). Of the at least 44 putative lysine acetyltransferases with recognisable KAT domains, we focused on the 23 KATs, which are expressed in most cell types and developmental stages and thus are likely to constitute the most relevant determinants of *Drosophila*'s histone acetylation system. This list includes (1) the extensively studied 'model' KATs (GCN5/PCAF, CBP/P300, MOF, HAT1, TIP60), (2) less well characterised KATs (KAT6 [MOZ/MORF], HBO1, ELP3, TAF1, ATAC2), (3) a mostly uncharacterised class of GCN5-related KATs (NAT6, NAT9, NAT10), (4) N-terminal acetyltransferases suggested to also acetylate internal lysines (NAA10, NAA20, NAA30, NAA40, NAA50, NAA60), (5) putative acetyltransferases with no recognizable direct homologue in non-*Drosophilid* species (CG5783, CG12560), (6) the acetyltransferase ECO and (7) a bi-functional enzyme containing a O-GlcNAcase - activity and potentially a KAT activity (MGEA5, also known as NCOAT or OGA).

The *D. melanogaster* genome encodes five recognisable zinc-dependent lysine deacetylases of the HDAC class (RPD3/HDAC1, HDAC3, HDAC4, HDAC6, HDAC11) and five NAD-dependent Sirtuins (SIR2/SIRT1, SIRT4, SIRT6, SIRT7). With the exception of HDAC11 and SIRT7, all KDACs are ubiquitously expressed and analysed in our study.

Because the majority of KATs and KDACs are not or only poorly characterised and consequently mutants and antibodies are lacking, we depleted KAT and KDAC proteins by RNA interference (RNAi) in the *Drosophila* model cell line KC. RNAi efficiently depletes proteins in *Drosophila* cell lines within 4 – 7 days (Zhou et al., 2014).

As a proof of concept we initially focused on the acetyltransferase MOF, for which well characterised mutants and antibodies exist. MOF levels were strongly diminished after incubating cells for 4 days with interfering RNAs directed against MOF transcript, and were undetectable after 5.5 days (Figures S4A, B). At this time H4.K16ac, the known product of MOF activity, was most strongly reduced (Figure S4C). The two different interfering RNAs yielded very similar results (Table S5). Because MOF is considered to be responsible for the majority of H4.K16ac in *Drosophila* and human cells (Smith et al., 2005; Conrad et al., 2012), we were surprised to observe that depletion of MOF in the female KC cell line reduced the amounts of H4.K16ac by only 40%.

To compare the acute RNAi depletion strategy with constitutive loss of gene function, we analysed the status of H4.K16ac in adult flies mutated for MOF (Gu et al., 1998). Similar to KC cells, *mof*² mutant females retained over 50% of H4.K16ac (Figure S4D). Apparently, other KATs besides MOF contribute to setting the H4.K16ac mark. Therefore, RNAi against MOF led to quantitatively similar reduction of H4.K16ac as loss-of-function mutation of the MOF gene. We conclude that RNAi is an appropriate method for our survey of KAT and KDAC activities.

We next ablated each KAT and KDAC in KC cells with two distinct interfering RNAs for 5.5 days and measured the abundance of 22 acetylation, 6 mixed acetylation-methylation and 11 methylation motifs as well as the 5 unmodified peptide backbones. Using RT-qPCR, we determined a mean knockdown efficiency of 80% across the 31 gene knockdowns (Figure S5). Multiple biological replicates increased the confidence to measure even subtle differences and demonstrated a high reproducibility among the RNAi pairs and across replicates (Table S5, Suppl. Note 3). We observed no major growth and cell cycle phenotypes upon depleting individual KATs and KDACs (data not shown), in agreement with a previous report (Kondo and Perrimon, 2011). Notable exceptions were cells in which CBP, TAF1 and RPD3 were targeted, which slowed their proliferation after four days of RNAi (data not shown).

The heat map in Figure 2A shows the relative changes of 24 histone acetylation motifs upon depletion of 23 KATs. Most KATs show a clear specificity for certain acetylation motifs, but CBP and NAA10 contribute to many motifs. KAT depletion leads to decreased levels of 104 acetylation motifs, with interesting specific effects. Many relations between KATs and particular histone motifs would have been missed had a conventional instead of a combinatorial analysis been applied (compare Figure 2B with 2A). Contrary to expectation, acetylation motifs did not only decline but at the same time the abundances of 160 motifs increased. Detailed analysis (see below) reveals that pre-existing acetylation and methylation patterns modulate the substrate selectivity for further modification.

Below, we use this dataset to i) describe the putative direct *in vivo* targets for KATs, ii) infer common properties of KATs as a class (Figure 3) and between the KATs and KDACs (Figure 4) and iii) describe the global changes of the histone acetylome as individual enzymes are removed from the

system (Figure 5). The analysis reports the response of the system to enzyme depletion, including primary effects reflecting direct enzyme-substrate relationship, and secondary effects for example of compensatory nature. To our knowledge, these 1364 individual enzyme - histone PTM relationships – monitoring the involvement of 23 KATs and 8 KDACs in 39 acetylation and methylation motifs – constitute the most comprehensive functional histone PTM dataset reported so far.

Mining the database for testable hypotheses and characteristics of acetyltransferase classes

An in-depth analysis of the data summarised in Figure 2 allows not only to shed light on controversial enzyme-substrate relationships but also to derive testable hypotheses about pathways that generate complex acetylation motifs. In Supplementary Note 4, this will be illustrated with two examples, the cytoplasmic HAT1 and with HBO1, substrates of which are currently debated (Suppl. Note 4 and Figures 3A-D).

Mining the data underlying Figure 2 demonstrates that many KATs specifically contribute to low abundant and combinatorial motifs. Most KATs have a narrow yet not absolute substrate specificity, which is modulated by adjacent acetylation (markedly observed for HAT1, HBO1, MOF, ATAC2, NAT10 and MGEA5) and methylation (KAT6, HBO1, NAT10 and others). This hitherto unappreciated level of crosstalk between neighbouring methylation and acetylation sites (see below) suggests that, like bromodomain recognition of histone tails, the interactions of KATs with substrate peptides are sensitive to the context of specific combinatorial motifs.

The realisation that most KATs have restricted substrate selectivity is in stark contrast to the widespread notion that most KATs are rather promiscuous enzymes (Sternier and Berger, 2000; Kouzarides, 2007; Lee and Workman, 2007). This view may have arisen by studies of the model acetyltransferase CBP, which, however, is a clear outlier within the KAT class (Figures 2, 3C, D). Loss of CBP leads to a reduction of many acetylation motifs, including acetylation at K5 and K8 of H4 and K14, K18, K23 and K27 on H3. Furthermore, whereas most KAT depletions reduce only a distinct subset of the motifs containing the putative primary target lysine, CBP RNAi reduces almost all H4 combinations containing acetylated K5 and K8 (Figure 3C). Evidently, CBP is a rather promiscuous KAT, which may explain its robust acetylation activity of many substrates *in vitro* (Sternier and Berger, 2000). Clearly, the enzymatic properties of CBP do not reflect common properties of KATs as a class.

Commonly, several KATs contribute to setting a histone acetylation motif. This fundamental conclusion echoes the combinatorial nature of histone modifications at the level of the enzymes that bring it about. While on average three KATs contribute to setting an acetylation motif, one often finds a dominant KAT with minor contributions by others. For example, CBP depletion leads to a reduction

of H4.K5acK8ac by 61% whereas MGEA5 and NAA10 each contribute only 30% and 23%, respectively (Figure 3E). Although HAT1 is the dominant KAT that sets H4.K5acK12ac, NAA10 and NAA60 also contribute substantially. Figure 3F plots the cumulative contributions of all KATs, calculated from the extent of motif reduction upon KAT depletion to all acetylation motifs. Notably, several bars approximate a cumulative depletion of 100% if all KAT RNAi effects are summed up (median = 113%, mean = 129%), suggesting that our systematic KAT depletion approach captures the majority of the cell's acetylation events. Cases where the cumulative contributions stay below 100% may be explained by functional redundancy between enzymes: the effect of an individual knockdown is underestimated because an alternative HAT takes over. For several rare and highly combinatorial motifs the cumulative contributions add up to more than 100%, which points to potential contributions of KDACs to the steady-state levels of acetylation motifs.

The relaxed specificity of deacetylases restricts the levels of rare, combinatorial motifs

The immediate reversibility that endows acetylation marks with regulatory potential requires strong deacetylase activities. Unlike KATs, KDACs are characterised by relaxed substrate specificity. Depletion of any of the eight commonly expressed KDACs increased the levels of many H4 and most H3 acetylation motifs in a similar way (Figure 4A). RPD3 is by far the dominant deacetylase in KC cells, and its depletion leads to strong gains of those rare, combinatorial motifs on H4 that bear regulatory potential.

Overall, motifs with multiple acetylation sites increase stronger than individually acetylated motifs. For example, di-acetylated H3.K9acK14ac increases more than 4 fold after ablation of RPD3. At the same time, the levels of the highly abundant H3.K14ac mark were only marginally increased by 20% whereas the ubiquitous H3.K23ac mark was not affected at all. We also did not detect any new acetylation of the lysines of the H3 and H4 tails that were not acetylated in unperturbed KC cells. In no case did the overall acetylation increase upon KDAC depletion lead to saturation and massive accumulation of even a simple motif (Figure 4B and Table S5).

The system's response to KAT deprivation effectively maintains global histone acetylation levels

Contrary to wide-held expectations we observed that the levels of many histone motifs *increased* if cells are deprived of histone acetyltransferases (Figure 2A). To explore this phenomenon more systematically, we summed up all gains and losses of the individual H3 and H4 acetylation motifs. To our surprise, we found that cells have balanced or even higher histone acetylation levels when depleted for most individual lysine acetyltransferases (Figures 5A and S6B). Ablation of only three KATs (KAT6, NAA10 and GCN5) lead to the expected significant reduction of global histone H3 and H4 acetylation. For a major group of 17 KATs, among them CBP and HBO1, the depletion did not change

the cell's total number of acetylated histone H3 and H4 peptides relative to non-acetylated peptides significantly, despite the fact that the abundance of some 'primary' acetylation motifs (the presumed direct products of KAT activity) were severely reduced. These losses were accompanied by gains in other, some times unrelated motifs (Figure 5B). We arrived at a similar conclusion when we counted the total number of *acetyl groups* instead of the number of *acetylated peptides* (Figure S6B).

Losses of acetylation are not only accompanied by gains within the same histone tail, as might have been expected, but also by gains on the other histone (Figures 2A, 5B). A prominent case is the depletion of CBP, where a reduction of the acetylation motifs from H4 (H4.K5ac and H4.K8ac) and H3 (H3.K14ac, H3.K18ac and H3.K18acK23ac) go along with an increase in H4.K12ac, H3.K9me2K14ac and H3.K23ac. Very often the acetylation gains are close to those sites that loose acetylation (Figures 2A, 5B), leading us to speculate that the system strives to compensate local charge imbalances upon KAT removal.

For a number of enzymes (NAT6, GC5783, NAA40 and TAF1) we do not detect any loss of acetylation upon depletion, but only gains (Figure 2A). These enzymes may not be KATs after all, or may strictly depend on cofactors that are not present in KC cells. Moreover, some KATs may only acetylate non-histone proteins (Sternier and Berger, 2000; Choudhary et al., 2009). Conceivably, the losses and gains in acetylation motifs upon perturbation of KATs reflect the response of a system that strives to compensate structural perturbations.

The methylation response to KAT depletion allows their functional classification

Depletion of most KATs lead to global changes in histone methylation marks. Hierarchical clustering of the methylation responses upon KAT deprivation revealed a clear grouping of KATs (Figure 6A). Many KAT depletions led to a substantial reduction of H3.K36me2/3, a mark associated with transcribed chromatin, as well as to a massive increase in the repressive H3.K27me3 mark, which indicates polycomb repression. This response is well expected for those KATs that are known transcription co-activators (GCN5, CBP, TIP60). The similarity of the response scored upon ablation of MGEA5 suggests that this poorly characterised KAT also function as a transcription activator. Interestingly, the putative acetyltransferases CG5783 and NAT6 are the only enzymes whose absence is accompanied by a substantial *increase* of K36me2. Similar, loss of NAA40 is accompanied with a severe reduction of the repressive K9me2 mark. These results suggest that CG5783, NAT6 and NAA40 may confer repressive functions.

Primary and secondary contributions of KATs to mixed acetylation/methylation motifs

For PTM motifs involving both, acetylation and methylation, it is difficult to conclude about the individual contribution of candidate enzymes. Figures 6B and 6C exemplify this challenge for KAT6 and HBO1.

Cells lacking KAT6 have reduced levels of H3.K14ac only in the context of tri-methylated lysine 9. Because H3.K9me3 alone is not reduced under these conditions, it is reasonable to assume that KAT6 directly acetylates K14ac on a tri-methylated K9 substrate. A similar argument can be made to conclude that KAT6 does not acetylate K14 if K9 is acetylated or mono-/di-methylated. The recognition of K14ac by antibodies may be modulated by adjacent methylation. By contrast, cells deprived of HBO1 not only reduce the H3.K9me3K14ac motif, but also the K9me3 mark alone, a secondary effect that may confuse the analysis of HBO1 substrates by other means.

Modulation of substrate specificity by molecular context

Acetyltransferases are commonly found in complexes with associated proteins, which may target the enzyme to specific chromosomal loci and fine-tune substrate selectivity. Variation in associated complex subunits may harness a KAT activity for cell-specific and developmental functions. This is well illustrated using the case of MOF, which is found in two distinct multi-protein complexes. In the context of the ubiquitous NSL (MBDR-2) complex MOF regulates housekeeping genes as a classical transcription co-activator (Feller et al., 2012; Lam et al., 2012). In male somatic cells, MOF additionally associates with the so-called male-specific-lethal (MSL) proteins to form the dosage compensation complex (MSL-DCC). The MSL-DCC compensates the X chromosome monosomy by an increased transcription output, which involves MOF-dependent H4.K16 acetylation of transcribed gene sequences (Prestel et al., 2010a).

In order to determine whether the protein complex environment modulates MOF's substrate selectivity we monitored changes in histone acetylation motifs after depleting male or female cells of MOF or of the diagnostic subunits of the NSL and MSL-DCC complex. RNA interference targeting MOF expression led to a reduction of H4.K16ac by about 40% in female cells or 60% in male cells (Figures 7A, S4D). MOF depletion in KC cells has to be seen against a robust increase of the almost three-fold more abundant H4.K12ac and additional gains in H3K9- and K14-containing acetylation motifs, leading to overall elevated levels of H3 and H4 acetylation. Likewise, H4.K16 loss in male S2 cells was accompanied by a clear increase in H4.K12ac and H4.K5ac. Ablation of the MSL-DCC subunit MSL1 in male cells essentially mirrored this effect (with less crosstalk to H4.K5ac), indicating that most MOF effects in S2 cells are mediated by the MSL-DCC. By comparison, interference with the NSL complex through ablation of its core subunit NSL1 only modestly affected H4.K16ac levels. Interestingly, we detected a strong decrease of H3.K27ac in both cell types and of H4.K5ac in female

cells lacking NSL1, but not MSL1 (Figure 7A, B). These findings illustrate the male-specific acetylation of H4.K16 over and above a lower, more general level. They also suggest that the NSL complex affects histone acetylation independent of MOF.

X chromosome dosage compensation is accompanied by global redistribution of acetylation and methylation marks

Due to the specific dosage compensation mechanism, male *Drosophila* cells contain more H4.K16ac than female cells (Figure 7C, D). In light of the observed global balancing of total acetylation levels (Figure 5) we wondered whether such chromosome-wide regulation had any consequences for the acetylation system as a whole. We, therefore, compared histone acetylation and methylation motifs in male and female cells as well as in male cells lacking MOF. Remarkably, male and female cells showed similar total histone acetylation levels, if all H3 and H4 acetylation motifs were summed up (Figure 7C). However, the distribution of acetyl groups between motifs in male cells differs dramatically from that of female cells, and the differences are due to MOF (Figures 7C, D, see also Table S4). Male cells displayed twice the amount of H4.K16ac, but reduced levels of several other acetylation motifs including H4.K5ac, H4.K12ac and H3.K14ac as compared to female cells. Conversely, depletion of MOF in male S2 cells reduced H4.K16ac and increases H4.K5ac, H4.K12ac and H3.K14ac to levels found in female cells.

Comparing methylation signatures between male and female flies indicate that dosage compensation correlates with an increased abundance of the repressive K27me3 mark as well as increased H3.K36me2/3 (Figure 7E). Abolishing dosage compensation in S2 cells by ablation of MOF converts the methylation pattern to a one that resembles females. Together, these results indicate that the evolution of dosage compensation not only boosted the levels of H4.K16 acetylation, but also triggered a network of secondary effects that may be interpreted as adaptations of the system to a global change in acetylation state.

The global response of the histone modification system upon MOF depletion is conserved between *Drosophila* and human cells

One of the surprise findings of our study is that a reduction of H4.K16 acetylation, which is thought to be a histone mark with unique structural implications, led to an inevitable increase of acetylation at the neighbouring K12. The robustness of the effect suggests K12ac may structurally compensate for a loss of K16ac. Since H4.K16 acetylation is thought to affect the basic properties of the chromatin fibre, such an effect should be observable in other species as well. To test this prediction we deprived human cells of MOF. In mammalian cells, MOF preferentially associates with subunits of the NSL complex,

but its substrate specificity is still under debate (Cai et al., 2010; Chelmicki et al., 2014). The ablation of MOF in HeLa cells not only reduced its supposed primary product, H4.K16ac, but also increased acetylation of the adjacent lysine 12, in striking resemblance to the effect in *Drosophila* (Figures 7F, S4F). Moreover, similar to female flies, where MOF functions as a global transcriptional co-regulator, human cells lacking MOF showed increased levels of H3.K9me2 and K27me3, while K9me3 and K36me3 are reduced (Figure 7G).

These results suggest that the global response of the system of histone modifications that we observed through our comprehensive, quantitative study of the acetylation motifs after depletion of critical enzymes, is a conserved feature of higher eukaryotic nuclei.

DISCUSSION

Although lysine acetylation was the first chemical modification described for histones already more than 50 years ago, our knowledge of the contributions of acetylation to the complex, combinatorial histone modification system is still superficial. Optimising LC-MS workflows and correcting for its inherent detection bias we generated a comprehensive inventory of all acetylation sites and many combinatorial motifs for histones H3 and H4 in an asynchronous population of *Drosophila* cells. Some highly abundant motifs may contribute to the general structure of chromatin, whereas rare, often combinatorial, motifs may transmit regulatory signals via bromodomain adaptor proteins.

Knowing the abundance of acetylation sites helps designing and interpreting genomic experiments, distinguishing rate-limiting from saturating components in a chromatin pathway and assessing the dynamic turnover of acetylation motifs (Waterborg, 2002; Zheng et al., 2013).

Acetyltransferases were the first histone modifying enzymes to be cloned, yet the substrate specificity of some KATs has remained controversial or even not known (Lee and Workman, 2007; Yuan and Marmorstein, 2013). Traditionally, identifying KAT substrates involved testing the acetylation status of selected lysines with antibodies, whose specificities have recently been questioned. Further problems arise from relating the apparent KAT substrate specificities determined *in vitro* to their physiological activities, which are commonly modulated by the molecular environment of the multi-subunit complexes they reside in. Remarkably, we found that most KATs display a narrow, yet not absolute, substrate specificity in cells, which is influenced by adjacent modifications. This contrasts common perceptions of rather promiscuous KAT activities that originate from the prominent characterisation of a few ‘model’ enzymes with relaxed substrate selectivity, such as CBP, which is a clear outlier of the class of KATs in this respect.

Our catalogue of the quantitative changes in histone acetylation and methylation motifs in response to ablating all known or suspected acetyltransferases and deacetylases allowed to re-evaluate

controversial enzyme-substrate relationships and to discover putative primary targets of uncharacterised KATs. This database can be further mined to derive testable hypotheses about the functional crosstalk between histone modifying enzymes and the assignment of KATs and KDACs to cellular pathways and disease states.

Surprisingly, we found that ablation of many KATs not only reduced acetylation of its putative primary target but also led to acetylation gains at secondary sites, such that the global level of histone acetylation in cells were maintained. Our survey of the secondary effects upon KAT depletion serves as point of entry to study the intertwined regulatory network of histone modifications as a *system*.

Related phenomena that suggest ‘compensation’ of a modification loss by corresponding gains on other proteins have occasionally been observed before. Analysing the phospho-proteome of a collection of yeast kinase and phosphatase mutants, Aebersold and colleagues reported a similar number of gains and losses in phosphorylation sites across all mutants together (Bodenmiller et al., 2010). Voss and colleagues found that reduced levels of H3.K14ac in *hbo1* knockout mice were accompanied by increases in H4.K16ac and to a lesser extent in H4.K5ac and H3.K9ac (Kueh et al., 2011). Ge and co-workers made similar findings for *cbp/p300* double deletion lines, which displayed increased levels of H3.K23ac and H3.K9me1K14ac (Jin et al., 2011). These related findings in several species may point to a fundamental, conserved principle.

Conceivably, the loss of a particular acetylation motif evokes a compensatory acetylation of a different, but functionally redundant motif by another KAT. The reciprocal responses between H4.K12ac and H4.K16ac (HBO1, MOF) and involving H4.K5 and H4.K8 (CBP) support such a view. However, particularly critical alterations in acetylation, such as losses in highly abundant, structural acetylation motifs, or of motifs with very specific functions during organismal development may not be compensated for (Voss and Thomas, 2009).

Considering histone modifying complexes as components of a system that satisfies local functional heterogeneity as well as global structural constraints, sudden perturbation may trigger feedback regulation to attenuate the impact. In this respect, global charge distribution and neutralisation comes to mind. Histone acetylation significantly diminishes the total number of positive charges in chromatin. We speculate that ablating KATs triggers a homeostatic re-adjustment of global charge to maintain the stability of chromatin according to polyelectrolyte theory (Clark and Kimura, 1990). In support of this hypothetical function, Kurdistani and colleagues reported a correlation between the global acetylation levels and the intracellular pH levels (McBrian et al., 2013).

The fact that lysines can be modified with several mutually exclusive chemical groups constitutes a basic principle of functional crosstalk between diverse modification-based signalling pathways. Further crosstalk is evident at the level of the modification enzymes, which are sensitive to pre-set modifications on histone tails. Monitoring lysine methylation and mixed acetylation-methylation

motifs enabled us to identify new candidates, where a pre-set methylation mark may promote (KAT6: H3.K9me3 → K14ac) or inhibit (NAA10: H3.K9me2 ↔ K14ac) acetylation. Moreover, we observed that ablating individual KATs led to dramatic reduction (GCN5: H3.K79me2) or increase (CBP: H3.K27me3) in methylation sites. Reciprocal correlations between the loss of histone lysine methyltransferases and altered histone acetylation have been documented before (Plazas-Mayorca et al., 2010; Sinha et al., 2010). Our study suggests that crosstalk among lysine acetylation and methylation is widespread and must be considered within the system of a comprehensive modification landscape.

The targeting and substrate selectivity of a KAT may be tuned by associated cofactors. Using the example of MOF, which resides in two distinct complexes with non-overlapping function (the MSL dosage compensation complex and the NSL transcription co-activator complex), we show that while the primary targets are similar across cell types and species (with notable quantitative differences), secondary effects are strongly influenced by the associated co-factors. Depletion of the NSL complex, but not the MSL complex, caused a massive reduction of the enhancer mark H3.K27ac, in interesting coincidence with the recent mapping of the human NSL complex to enhancers (Chelmicki et al., 2014). We cannot distinguish whether NSL subunits direct MOF's enzymatic activity towards K27 or whether the NSL complex regulates another KAT that acetylates H3.K27ac. It has also been suggested that MOF within the NSL complex acetylates K5 and K8 in addition to K16 (Cai et al., 2010), yet although we detect reduced amounts of H4.K5ac in female cells lacking NSL1, this is not observable for MOF depletions in S2, KC or HeLa cells.

Ablation of MOF activity in the context of the dosage compensation complex led to the expected reduction of H4.K16ac, but surprisingly, this was associated with a pronounced increase in the neighbouring H4.K12ac. Since the only known function of H4.K16 acetylation is to hinder the folding of the nucleosome fibre, we conclude that acetylation of lysine 12 can also fulfil this function to some extent. Interestingly, the depletion of the dosage compensation system further affected the recovery of distinct methylations. H3.K36me3, a mark that is placed co-transcriptionally and thus marks active genes was significantly reduced. Considering that the dosage-compensated chromatin amounts to less than 2% of the genome, our finding of massively increased, MSL-DCC-dependent bulk H3.K36me2/3 levels suggests that this mark almost saturates the nucleosomes along X chromosomal, dosage compensated genes. The high density may contribute to the observed genomic spreading capability of the MSL-DCC, which involves recognition of H3.K36me3 by MSL3 (Larschan et al., 2007).

Furthermore, depletion of MOF in male cells triggered a co-depletion of the repressive H3.K9me3 and H3.K27me3 levels, leading us to speculate that repressive chromatin attenuates the powerful activation potential of H4.K16 acetylation to arrive at the balanced levels that characterise dosage compensation (Prestel et al., 2010b). Our results define a point of entry into a quantitative assessment of the

epigenetic principles that enable the chromosome-wide transcriptional fine-tuning in vital processes, such as dosage compensation.

KAT and KDAC inhibitors are considered promising therapeutic agents in the combat against complex diseases including cancer and neurological disorders. Our work shows that depleting single KAT activities leads to complex alterations of the epigenome, of which we monitored the reduction of *bona fide* primary substrates, the global re-distribution of acetyl groups to secondary sites and changes to methylation of histones. Increased knowledge of the systemic response of the chromatin modification network will be required for a more targeted utilization of drugs in a clinical setting.

EXPERIMENTAL PROCEDURES

Cell Lines and RNAi

Cultivation of KC and S2 cells and RNAi were carried out essentially as described before (Feller et al., 2012). For detailed information, see Supplemental Experimental Procedures.

LC-MS for histone PTM quantification

Acid extracted, chemically acetylated and trypsinised histone peptides were separated on a HPLC C18 analytical column and electrosprayed into an LTQ-Orbitrap Classic. MS was operated in a targeted mode and data was analysed with Thermo Xcalibur, Excel and R.

Detailed information can be found in the Supplementary Notes 1 and 2 and in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, five Supplementary Notes, seven tables and Supplemental Experimental Procedures.

AUTHOR CONTRIBUTIONS

P.B.B. and C.F. conceived and designed the study. C.F. coordinated the project, performed most of the wet-laboratory experiments and analysed the data. I.F. and C.F. developed the LC-MS workflow. I.F. processed most samples for LC-MS and performed MS measurements. A.I. supervised LC-MS experiments and contributed important reagents. P.B.B. supervised the project. P.B.B. and C.F. wrote the manuscript and all authors revised the final version.

ACKNOWLEDGMENTS

We are grateful to the following colleagues: Dirk Schwarzer for providing synthetic peptides, Matthias Prestel and Heike Mittlöhner for help with fly genetics and maintenance, Thomas Blasi, Tobias Straub and Sascha Trostorff for advice on calculations and statistics, Ana Villar-Garea, Carsten Marr, Tobias Straub and Thomas Blasi for critical comments on the manuscript and members of the Becker lab for helpful discussions. This work was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n°293948. C.F. is a fellow of the International Max-Planck Research School (IMPRS-LS).

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD001394.

The authors declare no competing financial interests.

REFERENCES

- Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhhattar, R., *et al.* (2007). New nomenclature for chromatin-modifying enzymes. *Cell* *131*, 633-636.
- Bodenmiller, B., Wanka, S., Kraft, C., Urban, J., Campbell, D., Pedrioli, P.G., Gerrits, B., Picotti, P., Lam, H., Vitek, O., *et al.* (2010). Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Science signaling* *3*, rs4.
- Cai, Y., Jin, J., Swanson, S.K., Cole, M.D., Choi, S.H., Florens, L., Washburn, M.P., Conaway, J.W., and Conaway, R.C. (2010). Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. *The Journal of biological chemistry* *285*, 4268-4272.
- Chelmicki, T., Dundar, F., Turley, M.J., Khanam, T., Aktas, T., Ramirez, F., Gendrel, A.V., Wright, P.R., Videm, P., Backofen, R., *et al.* (2014). MOF-associated complexes ensure stem cell identity and Xist repression. *eLife* *3*, e02024.
- Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* *325*, 834-840.
- Clark, D.J., and Kimura, T. (1990). Electrostatic mechanism of chromatin folding. *Journal of molecular biology* *211*, 883-896.
- Conrad, T., Cavalli, F.M., Holz, H., Hallaceli, E., Kind, J., Ilik, I., Vaquerizas, J.M., Luscombe, N.M., and Akhtar, A. (2012). The MOF chromobarrel domain controls genome-wide H4K16 acetylation and spreading of the MSL complex. *Developmental cell* *22*, 610-624.
- Consortium, E.P., Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* *489*, 57-74.
- Cox, J., and Mann, M. (2011). Quantitative, high-resolution proteomics for data-driven systems biology. *Annual review of biochemistry* *80*, 273-299.
- Dawson, M.A., and Kouzarides, T. (2012). Cancer epigenetics: from mechanism to therapy. *Cell* *150*, 12-27.
- Feller, C., Prestel, M., Hartmann, H., Straub, T., Soding, J., and Becker, P.B. (2012). The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset. *Nucleic acids research* *40*, 1509-1522.
- Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Muller, S., Pawson, T., *et al.* (2012). Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* *149*, 214-231.
- Fischle, W., Wang, Y., and Allis, C.D. (2003). Histone and chromatin cross-talk. *Current opinion in cell biology* *15*, 172-183.

- Fuchs, S.M., Krajewski, K., Baker, R.W., Miller, V.L., and Strahl, B.D. (2011). Influence of combinatorial histone modifications on antibody and effector protein recognition. *Current biology : CB* 21, 53-58.
- Garcia, B.A., Mollah, S., Ueberheide, B.M., Busby, S.A., Muratore, T.L., Shabanowitz, J., and Hunt, D.F. (2007). Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nature protocols* 2, 933-938.
- Graff, J., and Tsai, L.H. (2013). Histone acetylation: molecular mnemonics on the chromatin. *Nature reviews Neuroscience* 14, 97-111.
- Gu, W., Szauter, P., and Lucchesi, J.C. (1998). Targeting of MOF, a putative histone acetyltransferase, to the X chromosome of *Drosophila melanogaster*. *Developmental genetics* 22, 56-64.
- Imhof, A., Yang, X.J., Ogryzko, V.V., Nakatani, Y., Wolffe, A.P., and Ge, H. (1997). Acetylation of general transcription factors by histone acetyltransferases. *Current biology : CB* 7, 689-692.
- Jin, Q., Yu, L.R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.E., Wang, C., Brindle, P.K., Dent, S.Y., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *The EMBO journal* 30, 249-262.
- Jung, H.R., Sidoli, S., Haldbø, S., Sprenger, R.R., Schwammle, V., Pasini, D., Helin, K., and Jensen, O.N. (2013). Precision Mapping of Coexisting Modifications in Histone H3 Tails from Embryonic Stem Cells by ETD-MS/MS. *Analytical chemistry* 85, 8232-8239.
- Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., *et al.* (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471, 480-485.
- Kondo, S., and Perrimon, N. (2011). A genome-wide RNAi screen identifies core components of the G(2)-M DNA damage checkpoint. *Science signaling* 4, rs1.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- Kueh, A.J., Dixon, M.P., Voss, A.K., and Thomas, T. (2011). HBO1 is required for H3K14 acetylation and normal transcriptional activity during embryonic development. *Molecular and cellular biology* 31, 845-860.
- Lam, K.C., Muhlpfordt, F., Vaquerizas, J.M., Raja, S.J., Holz, H., Luscombe, N.M., Manke, T., and Akhtar, A. (2012). The NSL complex regulates housekeeping genes in *Drosophila*. *PLoS genetics* 8, e1002736.
- Larschan, E., Alekseyenko, A.A., Gorchakov, A.A., Peng, S., Li, B., Yang, P., Workman, J.L., Park, P.J., and Kuroda, M.I. (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Molecular cell* 28, 121-133.
- Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. *Nature reviews Molecular cell biology* 8, 284-295.
- Lin, S., Wein, S., Gonzales-Cope, M., Otte, G.L., Yuan, Z.F., Afjehi-Sadat, L., Maile, T., Berger, S.L., Rush, J., Lill, J.R., *et al.* (2014). Stable Isotope labeled histone peptide library for histone post-

translational modification and variant quantification by mass spectrometry. *Molecular & cellular proteomics* : MCP.

Marx, H., Lemeer, S., Schliep, J.E., Matheron, L., Mohammed, S., Cox, J., Mann, M., Heck, A.J., and Kuster, B. (2013). A large synthetic peptide and phosphopeptide reference library for mass spectrometry-based proteomics. *Nature biotechnology* *31*, 557-564.

McBrian, M.A., Behbahan, I.S., Ferrari, R., Su, T., Huang, T.W., Li, K., Hong, C.S., Christofk, H.R., Vogelauer, M., Seligson, D.B., *et al.* (2013). Histone acetylation regulates intracellular pH. *Molecular cell* *49*, 310-321.

Parthun, M.R. (2012). Histone acetyltransferase 1: More than just an enzyme? *Biochimica et biophysica acta* *1819*, 256-263.

Pesavento, J.J., Bullock, C.R., LeDuc, R.D., Mizzen, C.A., and Kelleher, N.L. (2008). Combinatorial modification of human histone H4 quantitated by two-dimensional liquid chromatography coupled with top down mass spectrometry. *The Journal of biological chemistry* *283*, 14927-14937.

Plazas-Mayorca, M.D., Bloom, J.S., Zeissler, U., Leroy, G., Young, N.L., DiMaggio, P.A., Krugylak, L., Schneider, R., and Garcia, B.A. (2010). Quantitative proteomics reveals direct and indirect alterations in the histone code following methyltransferase knockdown. *Molecular bioSystems* *6*, 1719-1729.

Prestel, M., Feller, C., and Becker, P.B. (2010a). Dosage compensation and the global re-balancing of aneuploid genomes. *Genome biology* *11*, 216.

Prestel, M., Feller, C., Straub, T., Mitlohner, H., and Becker, P.B. (2010b). The activation potential of MOF is constrained for dosage compensation. *Molecular cell* *38*, 815-826.

Rothbart, S.B., Lin, S., Britton, L.M., Krajewski, K., Keogh, M.C., Garcia, B.A., and Strahl, B.D. (2012). Poly-acetylated chromatin signatures are preferred epitopes for site-specific histone H4 acetyl antibodies. *Scientific reports* *2*, 489.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nature reviews Molecular cell biology* *8*, 983-994.

Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callen, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I.A., *et al.* (2008). A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes & development* *22*, 2048-2061.

Sinha, I., Buchanan, L., Ronnerblad, M., Bonilla, C., Durand-Dubief, M., Shevchenko, A., Grunstein, M., Stewart, A.F., and Ekwall, K. (2010). Genome-wide mapping of histone modifications and mass spectrometry reveal H4 acetylation bias and H3K36 methylation at gene promoters in fission yeast. *Epigenomics* *2*, 377-393.

Smith, C.M., Gafken, P.R., Zhang, Z., Gottschling, D.E., Smith, J.B., and Smith, D.L. (2003). Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Analytical biochemistry* *316*, 23-33.

- Smith, E.R., Cayrou, C., Huang, R., Lane, W.S., Cote, J., and Lucchesi, J.C. (2005). A human protein complex homologous to the *Drosophila* MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Molecular and cellular biology* 25, 9175-9188.
- Sterner, D.E., and Berger, S.L. (2000). Acetylation of histones and transcription-related factors. *Microbiology and molecular biology reviews* : MMBR 64, 435-459.
- Suka, N., Suka, Y., Carmen, A.A., Wu, J., and Grunstein, M. (2001). Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Molecular cell* 8, 473-479.
- Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., *et al.* (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1016-1028.
- Voss, A.K., and Thomas, T. (2009). MYST family histone acetyltransferases take center stage in stem cells and development. *BioEssays : news and reviews in molecular, cellular and developmental biology* 31, 1050-1061.
- Waterborg, J.H. (2002). Dynamics of histone acetylation in vivo. A function for acetylation turnover? *Biochemistry and cell biology = Biochimie et biologie cellulaire* 80, 363-378.
- Young, N.L., DiMaggio, P.A., Plazas-Mayorca, M.D., Baliban, R.C., Floudas, C.A., and Garcia, B.A. (2009). High throughput characterization of combinatorial histone codes. *Molecular & cellular proteomics* : MCP 8, 2266-2284.
- Yuan, H., and Marmorstein, R. (2013). Histone acetyltransferases: Rising ancient counterparts to protein kinases. *Biopolymers* 99, 98-111.
- Zheng, Y., Thomas, P.M., and Kelleher, N.L. (2013). Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites. *Nature communications* 4, 2203.
- Zhou, R., Mohr, S., Hannon, G.J., and Perrimon, N. (2014). Inducing RNAi in *Drosophila* cells by soaking with dsRNA. *Cold Spring Harbor protocols* 2014.

FIGURE LEGENDS**Figure 1. Optimised liquid chromatography mass spectrometry (LC-MS) workflow for precise and accurate quantification of PTM motifs**

(A) Overview of the LC-MS workflow. D3AA method exemplified for H3 PTM isoform with four highly abundant modifications: H3.K9me3K14acK23acK27me2. (1) Chemical acetylation by d6-acetic anhydride transfers a deuterated (d3)-acetyl group to free lysines. (2) The subsequent trypsin digestion yields peptides of intermediate size. (3) Lysine methylation induces characteristic shifts in retention time, which allows quantification only based on MS1 spectra. Shown are ion intensity traces of the parent ions from the H3.K9-R17 peptide. To aid visualisation, intensities among retention time segments are scaled on the intensity of the third segment: Original intensities of the first and second segments are 6.3 and 3.1 fold higher. (4) The positional isomers with permutations of the lysine acetylation sites (H3.K9ac and H3.K14ac), require successive quantification using MS1 of double-charged parent ions (left) and MS2 of single-charged fragment ions (right).

(B) High precision quantification for 45 histone motifs in technical and biological replicate experiments. We determined the median coefficient of variations (CV) of 4.9% for three technical whole-workflow replicates using 2 million *Drosophila* KC cells (left), and CVs of 10.4% and 8.9% across five biological replicates for motifs which require successive MS1-MS2 or MS1-MS2-MS3 (centre, n=20) or MS1-only (right, n=25). The 45 histone motifs are shown in (D) and Table S4.

(C) LC-MS response correction factor improves accuracy for motifs containing lysine methylation sites. Synthetic peptides to acetylation and methylation modifications for H3_K9R17 peptide were individually chemically acetylated, trypsinised and measured by LC-MS. To derive the LC-MS response correction factor (right), raw values for the proteotypic peptide (left) were divided by raw values for the quantification tag (grey hexamer, centre) and signals for the H3.K9me1 peptide. Quantified synthetic peptides were obtained from JPT Peptide Technologies GmbH (Berlin).

(D) Abundance for histone PTM motifs after applying LC-MS response correction factor. Low (yellow), intermediate (orange) and high (red) abundance classes of histone acetylation motifs are indicated. Motifs containing no acetyl group are indicated in white. See Tables S4, S5 and Suppl. Note 1.3 for full dataset and variation estimates.

(E) Same as (D) for motifs where the LC-MS response correction factor had only intermediate confidence.

See Figures S1-S3.

Figure 2. Global and specific response to comprehensive perturbation of acetyltransferases

(A) Heat map displaying relative changes in abundance of 28 PTMs involving lysine acetylation in response to RNAi depletion of 23 KATs in KC cells. The first eight KATs are sorted according to the systematic listing by (Allis et al., 2007), the subsequent ones according to their sub-classes, including GCN5 related KATs (NATs) and N-terminal acetyltransferases (NAAs). Relative changes are normalised on control RNAi and log₂ scaled. Only significant changes are shown ($p < 0.05$, two-sided unpaired t test).

(B) Relative reductions of acetylation at individual lysines upon RNAi as in (A) cumulating all motifs containing the corresponding lysine. This illustrates the best-case outcome of a conventional analysis with hypothetical antibodies recognizing individual acetylated lysines with high specificity.

See Figures S5 and S6.

Figure 3. Most acetyltransferases have a narrow yet not absolute substrate specificity, which is modulated by the context of adjacent modifications

Statistical significance for all data in (A-F) was assessed with two-sided unpaired t test ($p < 0.05$) on log₂ (target RNAi/ctr RNAi).

(A) Quantification example of the histone H4 acetylation motifs for HBO1 RNAi. Statistically significant gains (red) and losses (blue) are indicated. Error bars display SEM (n=8).

(B) Scheme visualising the interdependence of combinatorial H4 motifs. HBO1-depleted cells display a highly selective reduction (blue arrows) of two acetylation motifs. Blue arrows point to significantly changed motifs (two-sided unpaired t test, $p < 0.05$).

(C,D) Quantification example and scheme as in (A) and (B), showing that CBP depletion reduces most H4 motifs containing acetylated lysine 5 or 8 (n=7).

(E) Cumulative contributions of KATs, which significantly (two-sided unpaired t test, $p < 0.05$) contribute to H4.K5acK8ac and H4.K5acK12ac.

(F) Cumulative contributions from all KAT depletion experiments reveals functional redundancy and points to the contribution of deacetylases. Cumulative KAT contribution was calculated from the extent of motif reduction upon KAT depletion filtered for statistically significant decreases (two-sided unpaired t test, $p < 0.05$).

Figure 4. KDACs generally show broad substrate specificity, and RPD3 dominates H4 de-acetylation

(A) KDACs display broad substrate specificity. Heat map showing relative changes of acetylation motif abundance in response to KDAC depletions, normalised to control RNAi and log₂ scaled. Only significant changes are shown ($p < 0.05$, two-sided unpaired t test on log₂ (KDAC/ctr)).

(B) Global increase of histone H3 and H4 acetylation upon depletion of any HDAC and Sirtuin. Relative changes of total H3, H4 and combined H3 and H4 acetylation levels normalised to control RNAi are shown. Significance was assessed on log₂ (KDAC/ctr) using two-sided unpaired t test ($p < 0.05$). RNAi to RPD3 and HDAC6 cause stronger gains in acetylation on H4 compared to H3. Error bars indicate SEM (for $n > 2$) or minimal and maximal value (for $n = 2$). See Table S5 for number of replicates ($n = 2 - 6$).

Figure 5. KAT deprivation induces and redistributes acetylation, which balances global histone acetylation levels

(A) KAT depletion may lead to globally reduced (blue), balanced (grey), or increased (red) acetylation levels. Shown is the global histone acetylation score, calculated by summing up all acetylation motifs on H3 and H4, and normalised to control RNAi. Statistically significant changes were assessed using a two-sided unpaired t test ($p < 0.05$) on log₂ (KAT/ctr). Error bars indicate SEM (for $n > 2$) or minimal and maximal value (for $n = 2$). See Table S5 for number of replicates ($n = 2 - 5$).

(B) Selective re-distribution of acetyl groups to specific secondary sites. Total loss (blue) and gain (red) of individual histone acetylation motifs are shown for cells deprived of CBP and HBO1. The contributions of the major motif changes of H3 or H4 acetylation are shown (filtered for significant changes, $p < 0.05$, two-sided unpaired t test).

See Figure S6.

Figure 6. Depletion of KATs and KDACs triggers a systemic alteration of the histone methylome

(A) Reorganisation of histone methylation sites after KAT deprivation. Heat map displays relative changes after KAT RNAi normalised to control RNAi. Only significant changes are shown ($p < 0.05$, two-sided unpaired t test.). Dendrograms were generated by unsupervised hierarchical clustering using the 'ward' algorithm on a Euclidean distance matrix. The H3.K27me₂ and H3.K79me₁ motifs did not change significantly in any KAT RNAi and therefore were not integrated in the heat map.

(B-C) Comparative analysis of acetylation-only, methylation-only and mixed acetylation-methylation motifs facilitates prioritization of putative KAT targets and identifies acetylation-methylation

crosstalk. Significantly reduced (blue) and increased (red) motifs are indicated (two-sided t test, $p < 0.05$). Error bars indicate SEM ($n=3$ for KAT6, $n=4$ for HBO1).

Figure 7: X-chromosome dosage compensation is accompanied by global redistribution of acetylation and methylation marks

(A) The MOF-containing MSL-DCC complex dominates the NSL complex for acetylation of H4.K16. Female KC or male S2 cells were depleted of MOF, MSL1 or NSL1, and the relative abundance of the indicated H4 acetylation (red box) was expressed relative to control cells that were treated with interfering RNAs directed against GST and GFP.

(B) Similar type of experiment as in (A), probing the effect of NSL1, MSL1 or MOF depletion on H3.K27ac levels.

(C) Global histone acetylation levels are similar between male S2 and female KC cells, but the distribution of individual sites differs. Cumulative bar plots summing up total levels of individual histone motifs.

(D) Similar experiment as in (A), revealing the re-distribution of acetyl groups upon MOF depletion in S2 cells

(E) Similar experiment as in (A), revealing that MOF depletion correlates with increased H3.K9me3, H3.K27me3 and H3 K36me2/3. Error bars in A, B, D and E indicate SEM ($n=3$ for NSL1 in S2 and Kc and MOF in S2; $n=4$ for MOF in Kc (for H3) and $n=11$ for MOF in Kc (on H4) and minimal/maximal values for MSL1 ($n=2$).

(F) Similar experiment as in (D), except that MOF was ablated in Hela cells.

(G) Similar experiment as in (E), except that MOF was ablated in Hela cells. Error bars in F and G represent SEM ($n=4$).

Statistical significance between control – target RNAi and S2 – KC are indicated by asterisks (two-sided t test, $p < 0.05$).

See Figure S4.

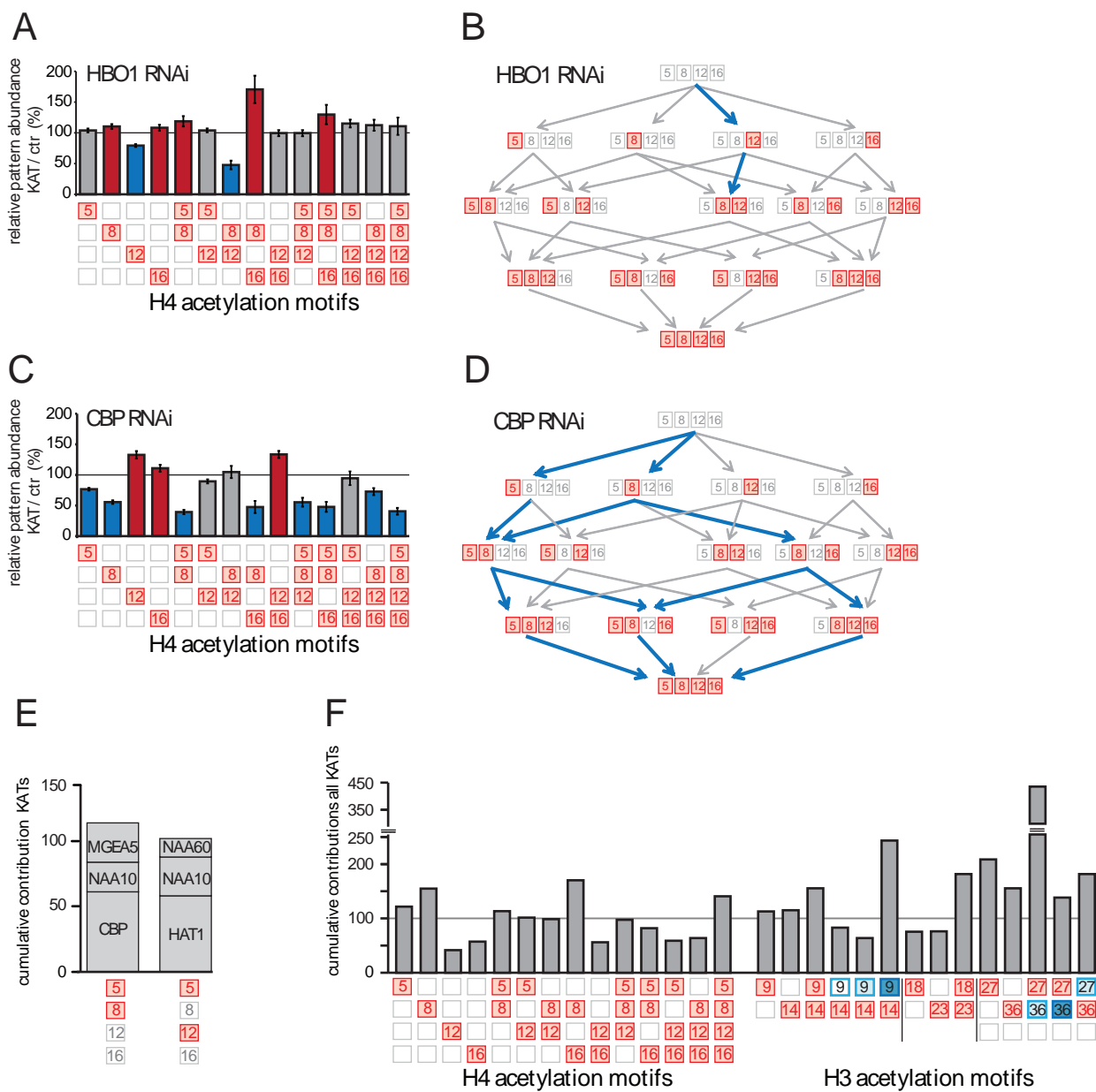


Figure 3

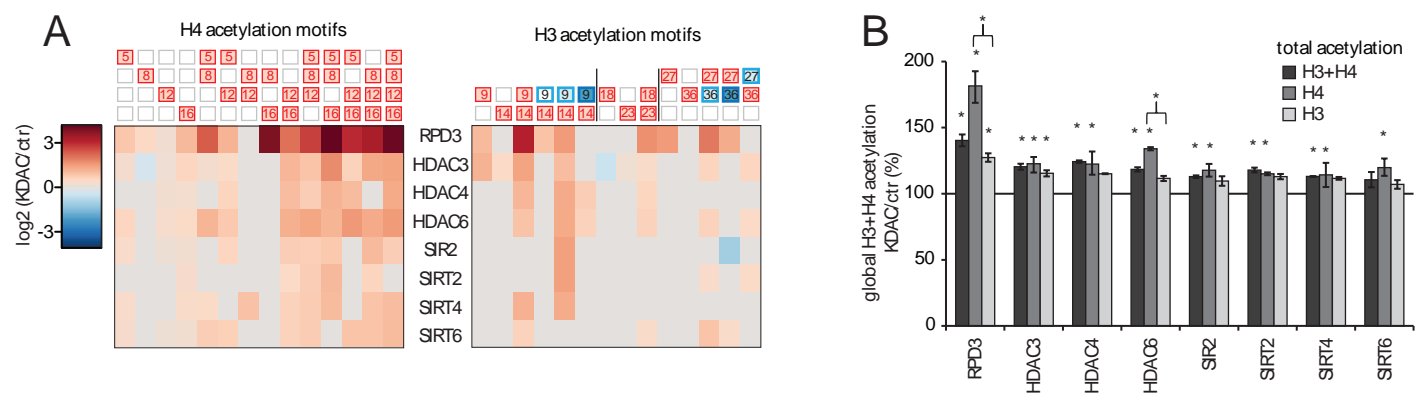


Figure 4

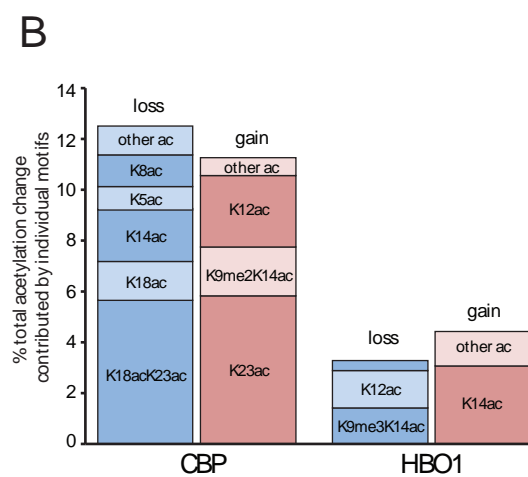
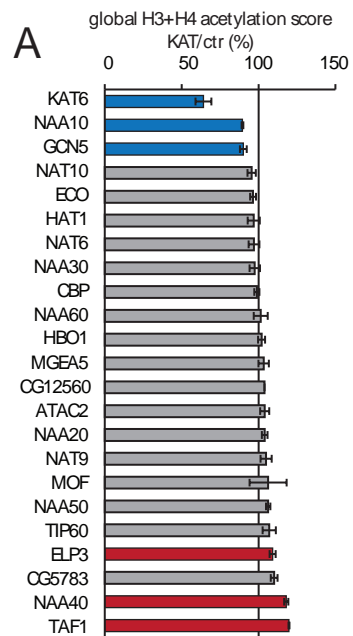


Figure 5

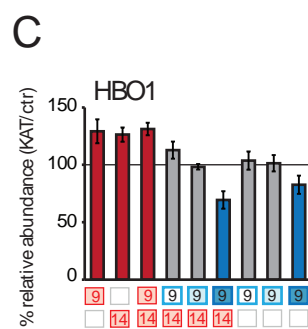
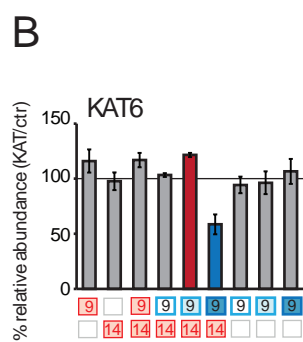
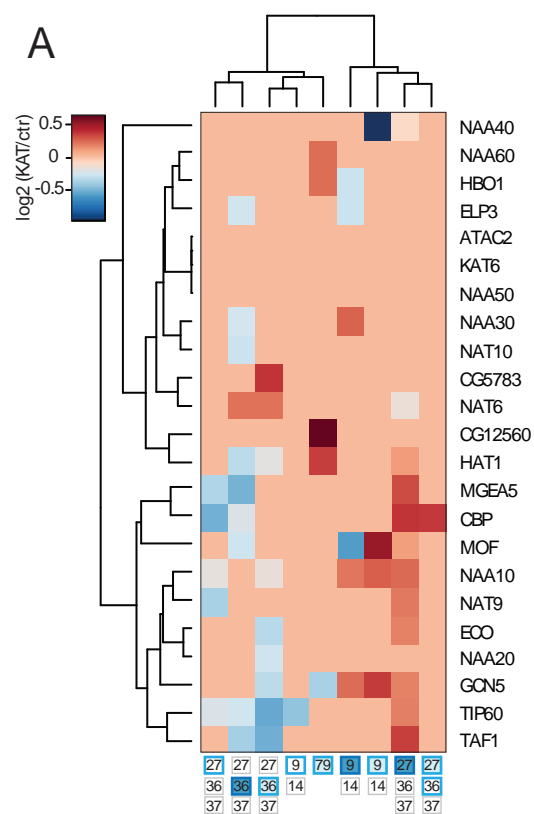


Figure 6

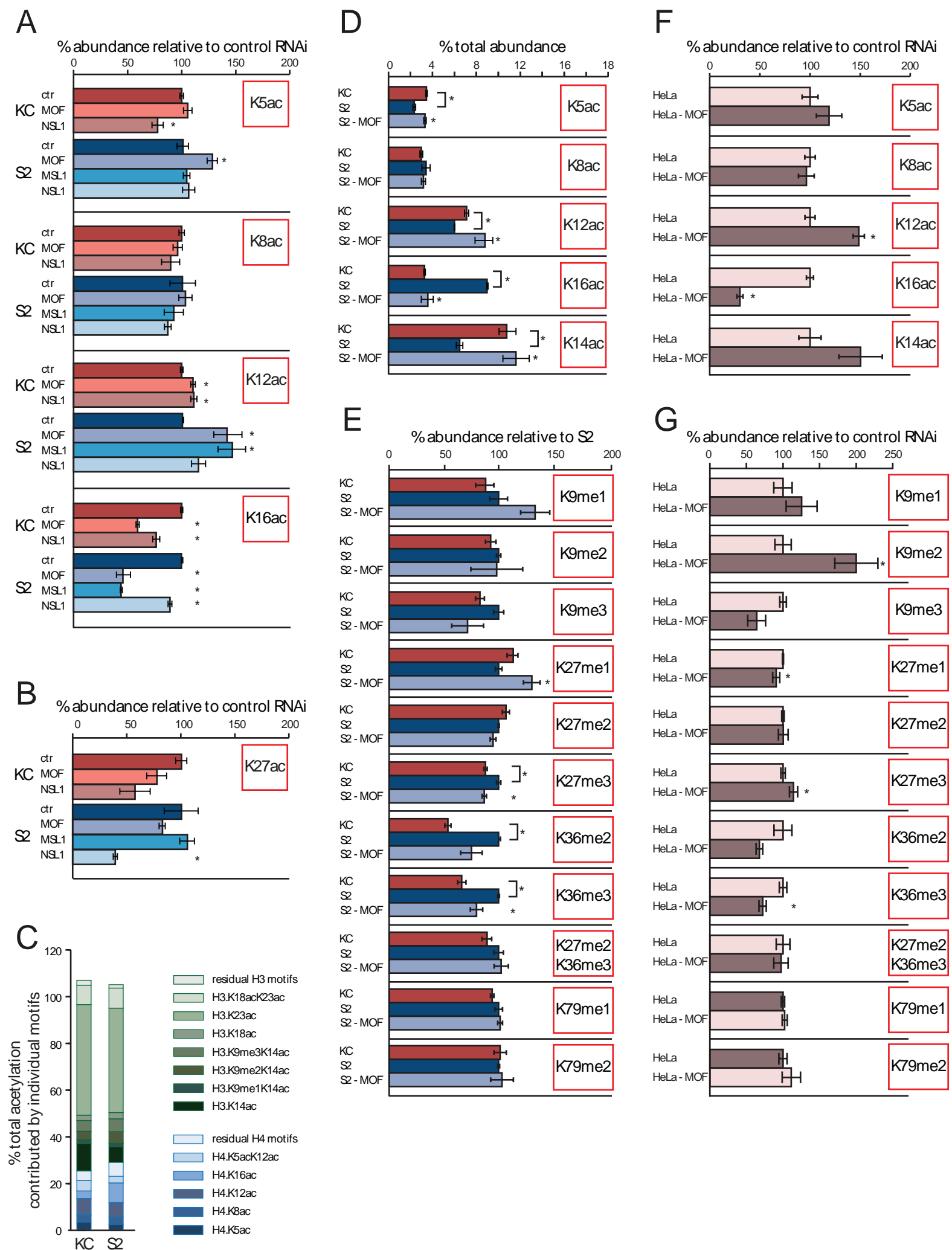


Figure 7

3.4.3 Supplementary data, tables and figures

Table of Contents

Supplementary Figure 1: Chromatographic separation of histone PTM motifs

Supplementary Figure 2: MS1-MS2-MS3 analysis for the quantification of isobaric positional isomers to H4.G4-R17

Supplementary Figure 3: MS1 and MS2 spectra for the quantification of H3 motifs

Supplementary Figure 4: Cells lacking MOF still retain substantial levels of H4.K16ac

Supplementary Figure 5: Quantification of mRNA levels after KAT and KDAC RNAi

Supplementary Figure 6: Most relative changes of acetylation motifs in response to KAT and KDAC ablation are not affected when correcting for the LC-MS bias

Supplementary Note 1: LC-MS workflow

Supplementary Note 2: LC-MS response correction factor increases accuracy for histone PTM quantification

Supplementary Note 3: Experimental design and statistical analysis

Supplementary Note 4: Mining the database to derive testable hypotheses and to shed light on controversial enzyme-substrate relationships

Supplementary Note 5: Homeostatic control of acetylation levels upon deprivation of KATs

Table S1: List MS quantifier

Table S2: Calculations to derive motif abundance from MS quantifier

Table S3: LC-MS response correction factor

Table S4: Histone PTM inventory for KC and S2 cells

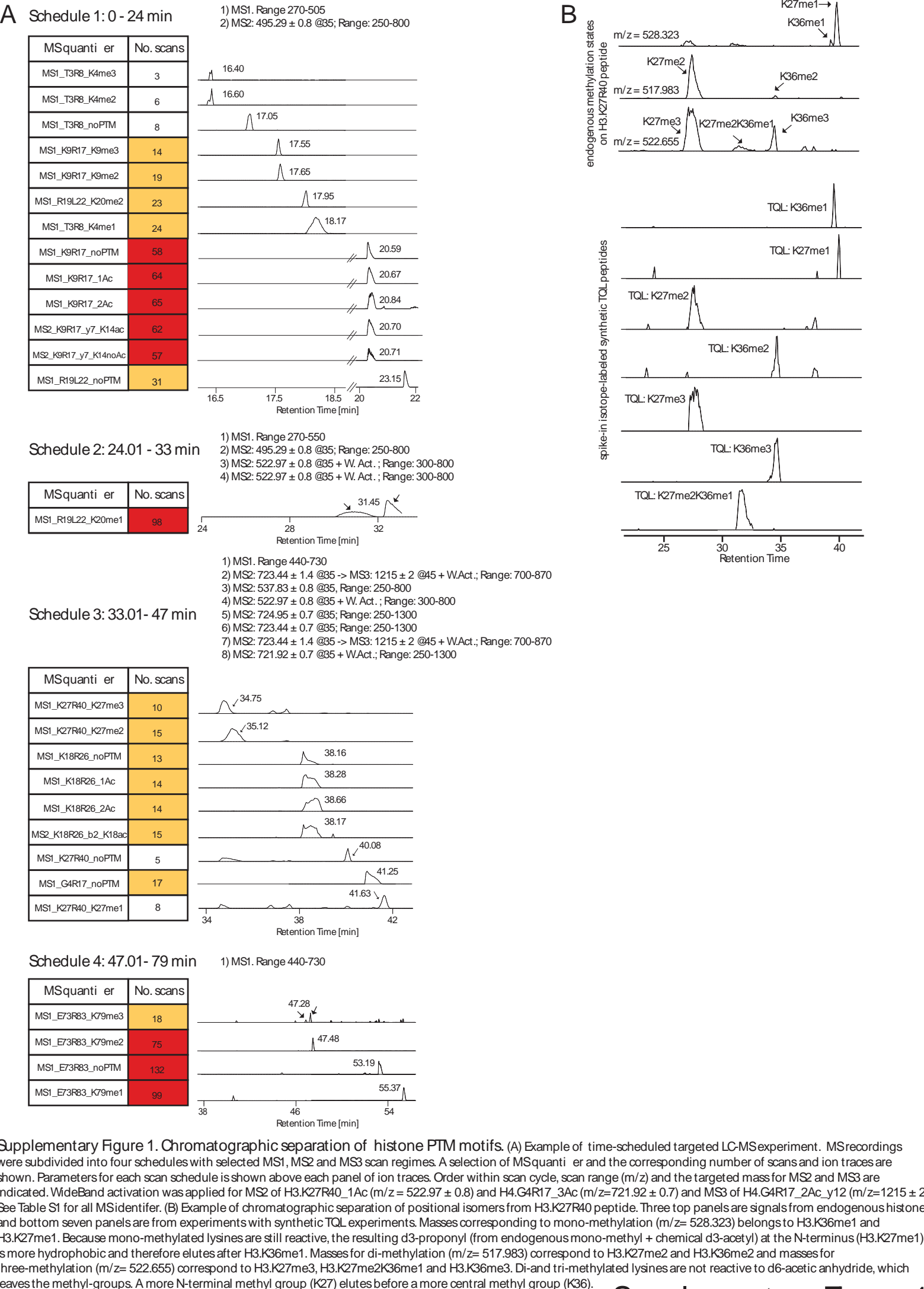
Table S5: Histone PTM inventory after KAT and KDAC RNAi in KC (accompanying Excel spreadsheet)

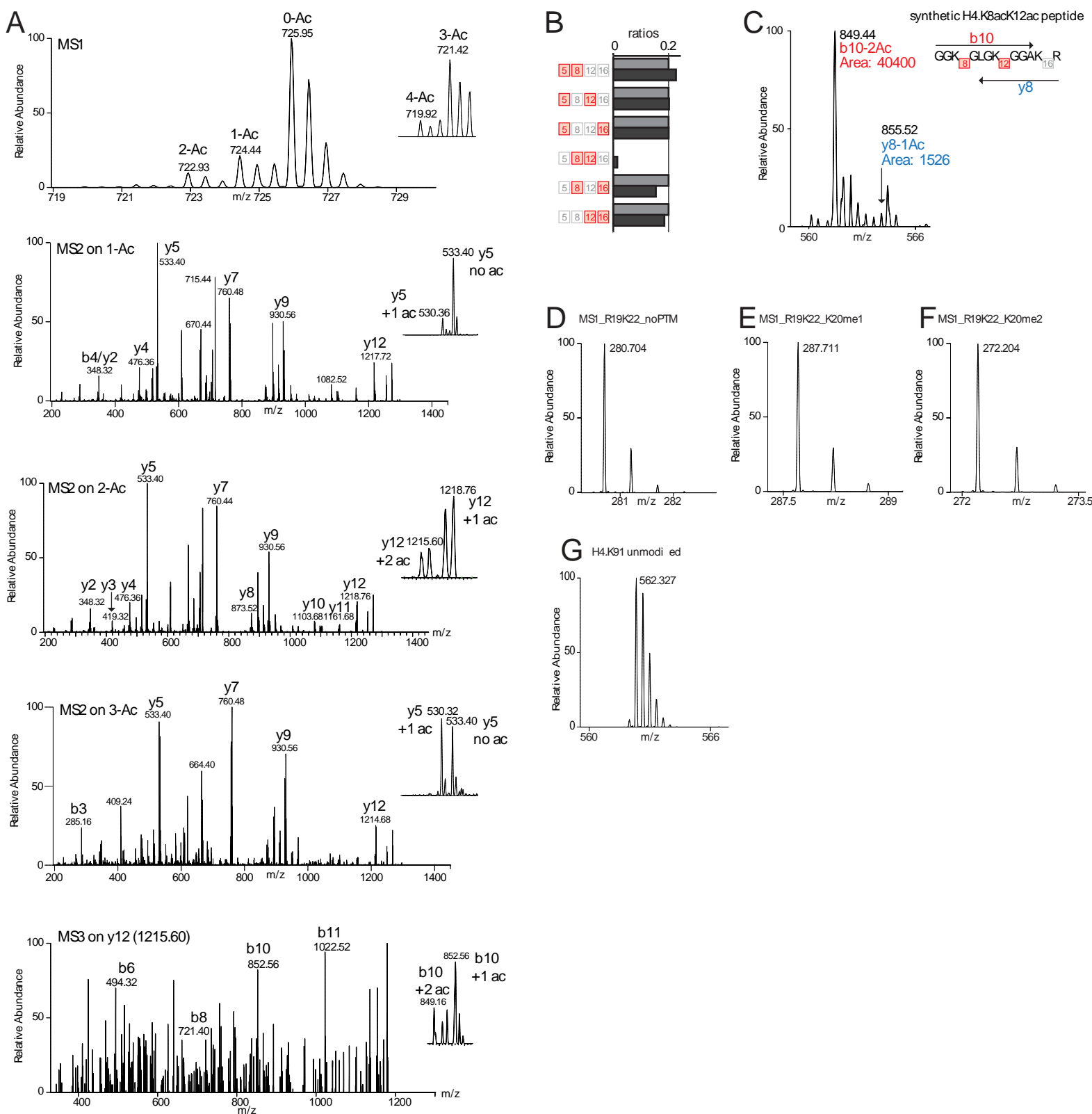
Table S6: List of *Drosophila melanogaster* KATs and KDACs

Table S7: List of primers used in this study

Supplemental Experimental Procedures

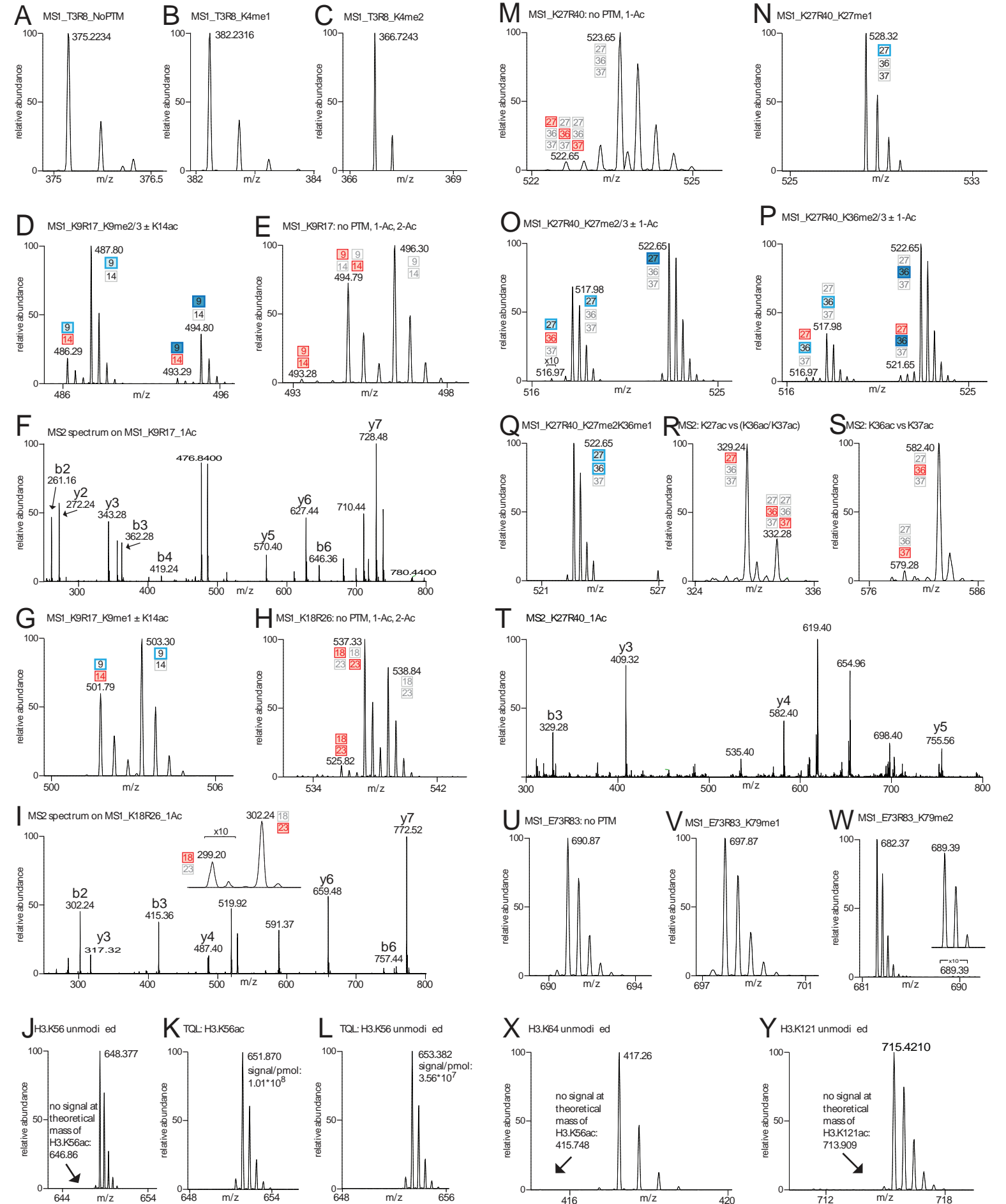
Supplemental References



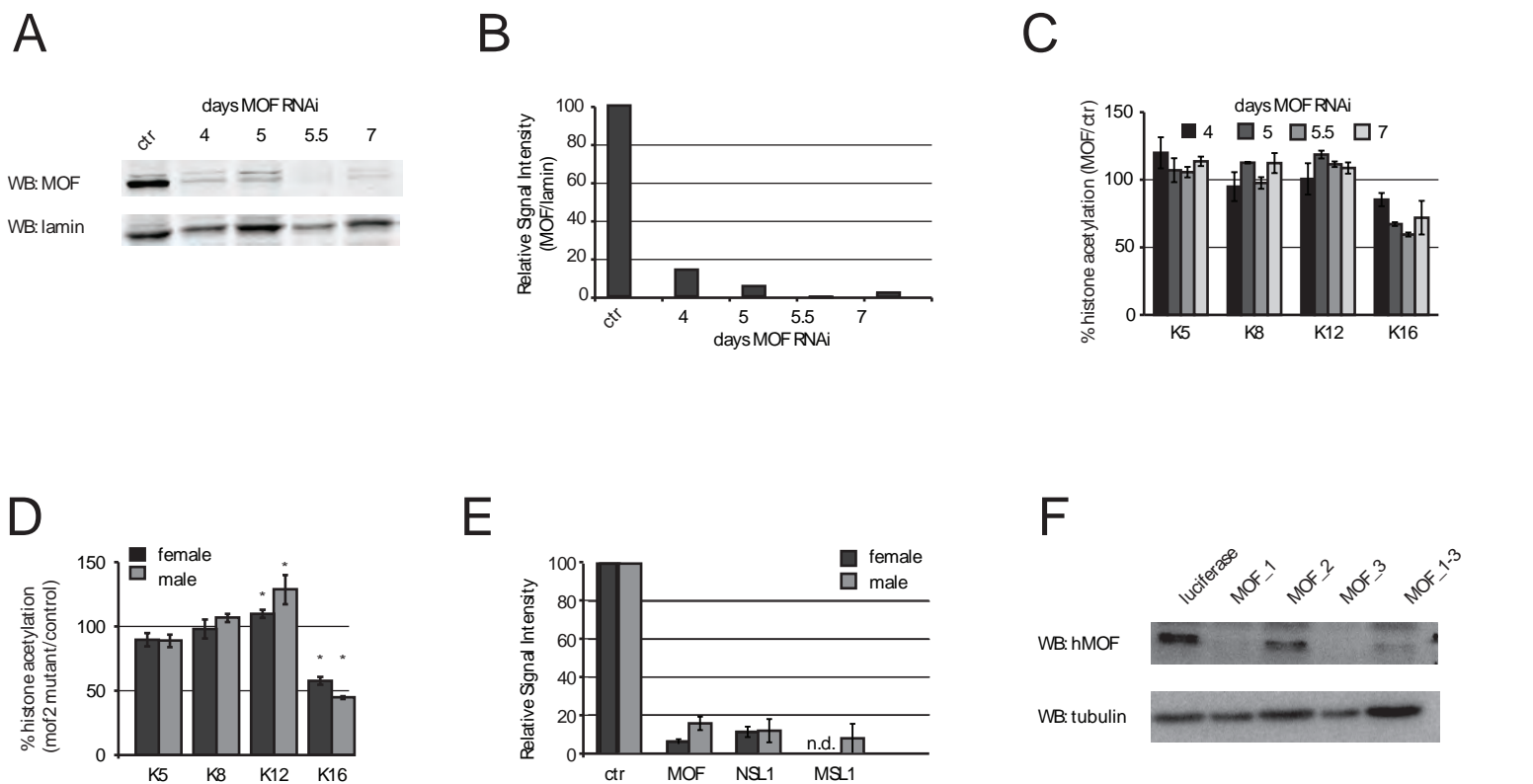


Supplementary Figure 2. MS1-MS2-MS3 analysis for the quantification of isobaric positional isomers to H4.G4-R17.

(A) Spectra for MS1, MS2 and MS3 of the H4.G4-R17 peptide. Note the shift of the MS3 spectrum by -18 caused by the loss of a water molecule. MS3 signals were quantified with b10 fragment ion pairs. Inserts for selected spectra ranges are shown (right). (B) Successive MS1-MS2-MS3 based quantification recovers accurate ratios in experiments with defined concentrations of synthetic peptides. Shown is a representative experiment where five out of six di-acetylated H4 motifs were mixed in equal ratios, processed and quantified via successive MS1-MS2-MS3. Measured ratios (black) are close to expected ratios (grey). (C) The b10 MS3 fragment ions can be used for quantification. The theoretical masses of the b10 MS3 fragment ions (849.52 if di-acetylated and 852.52 if mono-acetylated) are close to the theoretical masses of the y8 ions (852.56 if di-acetylated and 855.56 if mono-acetylated) yet y8 is not observed or only with a minor signals. The MS3 spectrum from an H4.K8ack12ac peptide (gift of D. Schwarzer, Tübingen) generates a dominant b10-2Ac ion at 849.44 (Area: 40400) yet only a minor y8-1Ac signal at 852.52 (Area: 1526), demonstrating that any interference from y8 is minimal, and the b10 ion pair can be used for quantification. (D-F) MS1 spectrum for H4.K20_noPTM (D), H4.K20me1 (E) and H4.K20me2 (F). (G) MS1 spectrum shows triple-charged unmodified H4.K79-R92 peptide, yet no signals for acetylated isoform (theoretical mass of H4.K91ac: 561.220).

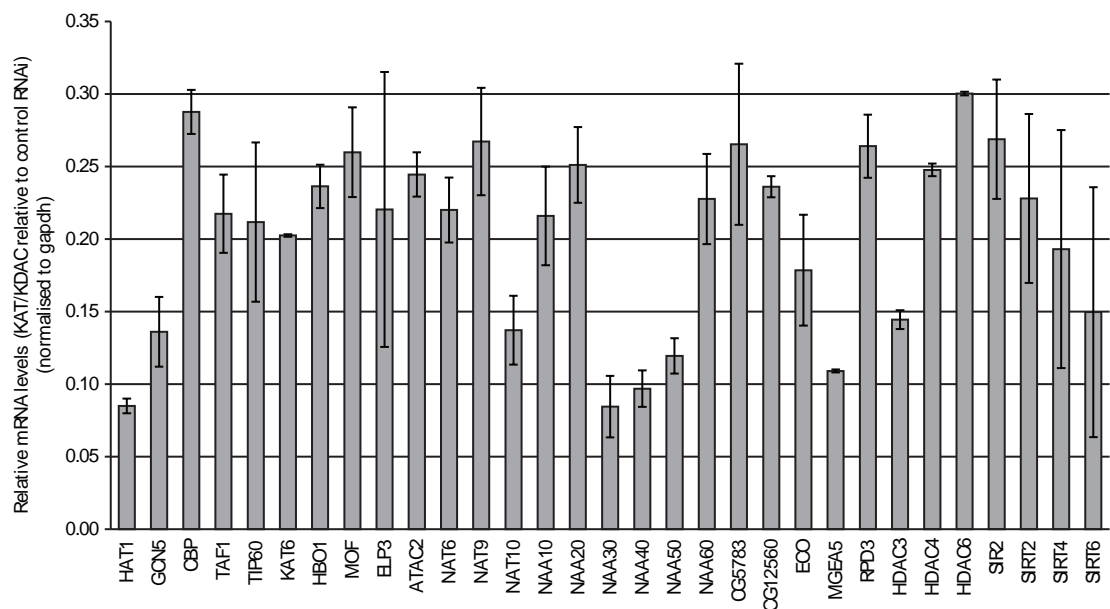


Supplementary Figure 3. MS1 and MS2 spectra for the quantification of H3 motifs.
 (A-C) MS1 spectra of H3.T3R-8 peptide, with no modification (A), H3.K4me1 (B) and H3.K4me2 (C). (D, E, G) MS1 spectra of H3.K9-R17 peptide, with H3.K9me2/3 ± K14ac (D), unmodified (E), mono- and di-acetylated H3.K9-R17 peptides (E) and H3.K9me1 ± K14ac (G). (F) MS2 on mono-acetylated H3.K9-R17 is required to distinguish H3.K9ac and H3.K14ac. Insert to diagnostic y7 ion pairs is shown in Figure 1A (panel 4). (H) MS1 spectrum of H3.K18-R26 peptide. (I) MS2 on mono-acetylated H3.K18-R26 is required to distinguish H3.K18ac and H3.K23ac. Insert shows diagnostic b2 ion pair. Note that signals around 299.20 were 10x enhanced. (J-L) Spectra for H3.Y54-R63 peptide. Robust identification of unmodified H3.K56 but not acetylated H3.K56ac in histones from KC cells (J), although synthetic TQL peptides to H3.K56ac (K) produce at least similar signals to unmodified H3.K56 TQL peptides (L). (M-Q) MS1 spectra of H3.K27-R40 peptide, with unmodified (M), H3.K27me1 (N), H3.K27me2 ± K36ac and H3.K27me3 (O), H3.K36me2/3 ± K27ac (P), H3.K27me2K36me1 (Q). (R-T) MS2 on mono-acetylated H3.K27-R40 is required to distinguish H3.K27ac, H3.K36ac and H3.K37ac. Full MS2 spectrum (T) and diagnostic b3 (R) and y11 (S) ion pairs are shown. (U-W) MS1 spectra of H3.E73-R83 peptide, with unmodified (U), mono-methylated (V) and di-methylated (W) H3.K79. The insert in W shows minute signals for H3.K79me3 that elutes shortly before H3.K79me2. (X, Y) Robust signals for unmodified H3.K64 and H3.K121, but no signal for acetylated isoforms are detected.

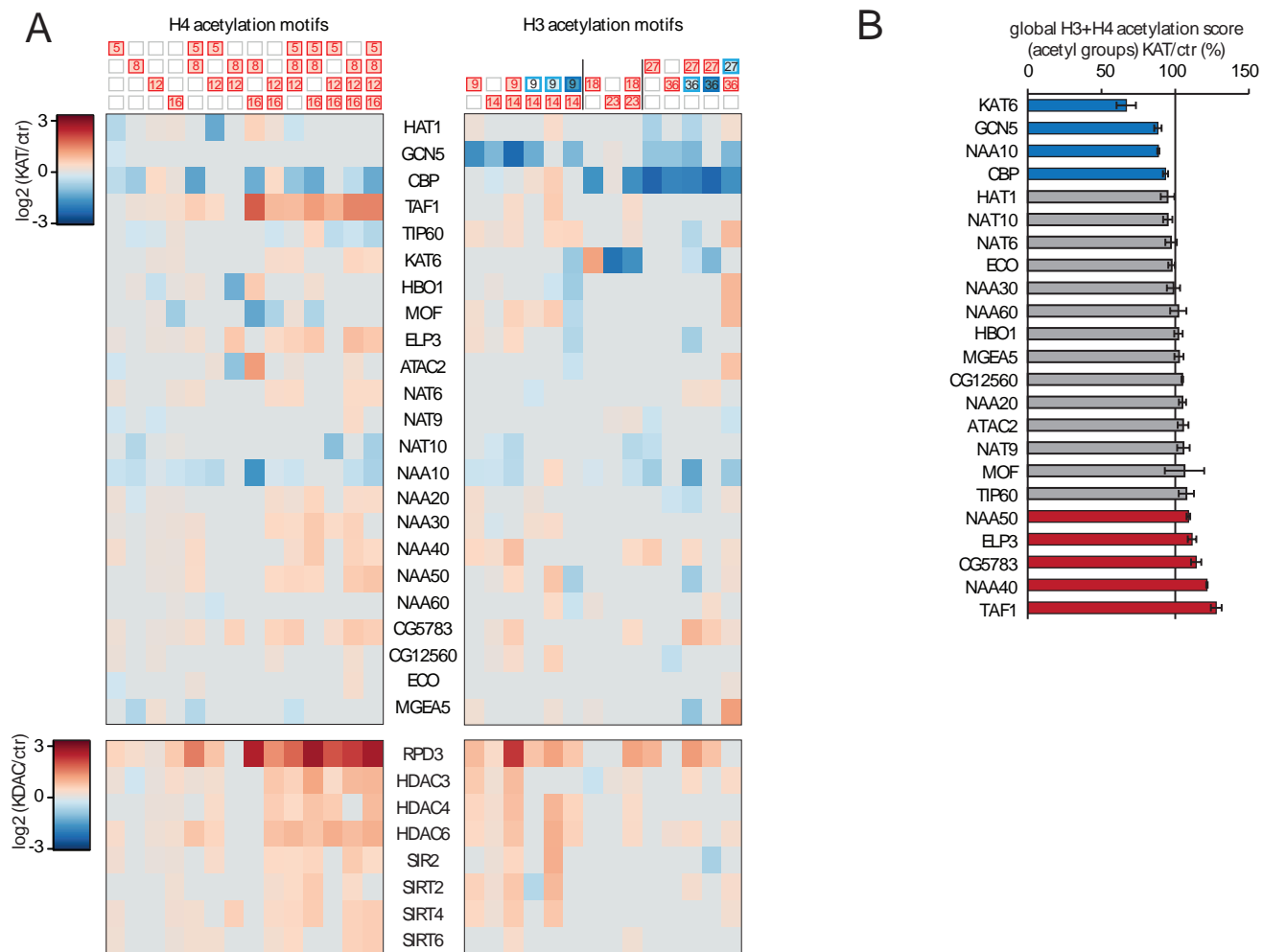


Supplementary Figure 4: Cells lacking MOF still retain substantial levels of H4.K16ac.

(A) Representative example of quantitative immuno blot analysis probing for MOF protein levels during a four to seven days time-course RNAi experiment in female Kc cells. Lamin protein levels served as a loading control. (B) Quantification of immuno blot signals from (A) using LI-COR (Biosciences). (C) Quantitative LC-MS analysis of mono-acetylated H4 motifs during a time-course RNAi experiment using two independent MOF RNAi constructs. Values were normalized to control RNAi. Error bars show maximum and minimum from two different RNAi constructs targeting MOF. (D) Quantitative LC-MS analysis of mono-acetylated H4 motifs in adult female (black) and 3rd instar larvae male (grey) mof2 mutants. Error bars represent SEM from 8 (female) and 3 (male) independent biological replicates. The increase of H4.K12ac and the decrease of H4.K16ac is statistically significant ($p < 0.05$ using unpaired two-sided t test). (E) Quantification of immuno blot analysis targeting MOF, NSL1 (NSL complex) and MSL1 (MSL-DCC complex) in Kc and S2 cells after 5.5 days of RNAi using LI-COR (Biosciences). Error bars represent SEM of at least three biological replicates. (F) ECL-based immuno blot analysis of RNAi experiment to target hMOF in HeLa cells for 2.5 days using three different siRNA constructs (lanes 2-4), a pool of 3 MOF siRNA constructs (lane 5) and a luciferase sequence as control.



Supplementary Figure 5: Quantification of mRNA levels after KAT and KDAC RNAi. Bar plots show results of RT-qPCR measurements after RNAi of indicated KATs and KDACs, normalised to control RNAi and gapdh. Error bars show minimal/maximal value from two different dsRNA RNAi constructs targeting the same KAT/KDAC gene (see Table S7 for RNAi and RT-PCR primers). In cases where only a single dsRNA RNAi construct was used, the replicate experiment was performed on another day.



Supplementary Figure 6: Most relative changes of acetylation motifs in response to KAT and KDAC ablation are not affected when correcting for the LC-MS bias

A) Heat map displaying relative changes in abundance of 28 histone acetylation motifs in response to KAT and KDAC ablation without applying the LC-MS response correction factor. Only significant changes are shown ($p < 0.05$, two-sided unpaired t test on log₂ ratios).

B) Global acetylation score calculated by counting acetyl groups instead of acetylation motifs. Here, the abundances of motifs with two acetyl-groups are multiplied by two, motifs with three acetyl-groups are multiplied by three, etc.

SUPPLEMENTARY NOTES

Supplementary Note 1: LC-MS workflow

1.1 Improved sensitivity, precision and accuracy to quantify positional isomers containing lysine acetylation

We streamlined the histone extraction protocol to enable rapid processing of many samples (see Supplemental Experimental Procedures). The low requirement of starting material (typically 1 - 5 x 10⁶ KC cells) allowed a cost-effective RNAi strategy and histone preparation without chromatographic enrichment. The direct extraction of histones from cells, rather than from nuclei or chromatin reduced preparation time and did not require usage of deacetylase inhibitors. Commonly used deacetylase inhibitors such as sodium butyrate alter the native histone acetylation pattern by inhibiting different deacetylases of the HDAC class with different efficiencies (Thorne et al., 1990; Fraga et al., 2005; Karmodiya et al., 2012). Our procedure aims at minimising the risk of altering the native histone acetylation pattern during sample preparation.

Histones were chemically acetylated using d6-deuterated acetic anhydride rather than propionylated to harmonise the physiochemical properties between the endogenously acetylated and non-acetylated lysines. This ensures similar properties during the LC-MS workflow, including comparable interactions during pre-LC (similar hydrophobicity, tube interactions etc.), LC (positional isomers co-elute and hence are electrosprayed at the same acetonitrile concentration) and MS (similar 'matrix effects' during ionization, fragmentation, MS acquisition), which results in improved accuracy among positional isomers. For example, using equimolar solutions of positional isomers from di-acetylated H4.G4-R17 peptides resulted in very similar ratios (Figure S2B). Furthermore, positional isomers for H3 peptides differing only in their site of acetylation display very similar LC-MS response correction factors (Table S3).

We developed a targeted MS protocol on a standard hybrid linear ion trap Orbitrap instrument (LTQ Orbitrap Classic) but the MS strategy can be implemented on different mass spectrometers capable of MS3 fragmentation. The general rationale is to iteratively acquire full or targeted survey scans (MS1) combined with targeted MS2 or MS3 spectra during the entire peptide elution or scheduled for pre-defined retention time windows (see Figure S1A for an example of scheduled RT windows). The scheduled spectra were recorded using a reduced window (e.g. MS1 acquisition window from 270 to 505 during the first retention time schedule) and less MS2/MS3 were scheduled. This improved the quantification of the MS identifier eluting within narrow LC peaks (e.g. H3.K9me2/3) through achieving a higher number of scans. We noted, however, that a few MS identifiers showed considerable shifts in their retention time within consecutive LC runs (in particular the non-methylated H3.K9R17 identifier with more than 5 minutes shift). Note that most of the data presented in the

current study was therefore acquired without retention-time scheduling. In contrast to previously developed targeted MS strategies (Schmidt et al., 2008; Gallien et al., 2012a; Gallien et al., 2012b; Peterson et al., 2012; Jaffe et al., 2013; Tang et al., 2014) our emphasis was on the high-precision quantification of combinatorial PTM motifs. To increase the MS duty cycle, MS1 scans were recorded within the Orbitrap at low resolution (7500 at 400 m/z) after accumulating 100,000 ions over a maximum time of 500 ms. To increase the precision and sensitivity for MS2 and MS3, enhanced zoom scans were acquired within the ion trap targeting the first, second and third isotope, setting a target value to 50,000 and accumulating for a maximal time of 100 ms. WideBand activation was performed to improve fragmentation for MS2 of the H3K27R40_1Ac ($m/z = 522.97 \pm 0.8$) and H4G4R17_3Ac ($m/z = 721.92 \pm 0.7$) identifier and for MS3 of the H4G4R17_2Ac_y12 ($m/z = 1215 \pm 2$) identifier. The targeted strategy enhanced the precision, reduced the interference and yielded a high density of scans.

To determine how many cells are required for high-precision analysis, we collected three times 1, 2 and 4 $\times 10^6$ cells for a LC-MS dose-response experiment. The median coefficient of variation (CV) across 45 histone motifs was 9% for the 1 million starting cell sample and around 5% for the 2 and 4 million cell samples (Table S4 and data not shown). We therefore used 2 to 4 million cells for most experiments. Notably, although we obtained sufficient MS signals and reliable motif ratios even with 2 $\times 10^5$ KC cell equivalents and 2 $\times 10^4$ mammalian cell equivalents if diluted down from cell suspensions, we did not succeed to robustly start with less than 1 $\times 10^6$ KC cells due to varying sample loss during various centrifugation steps and acid extraction.

1.2 Successive MS1-MS2-MS3 to quantify combinatorial motifs of the di-acetylated histone H4 positional isomers

The abundance of peptides with two out of four acetylations cannot be resolved with MS1 and MS2 analysis alone, because the set of linear equations for these motifs cannot be mathematically solved (as noted before by Phanstiel et al. 2008). One possible solution is to separate these positional isomers by chromatography (Young et al., 2009; Leroy et al., 2012). Potential drawbacks of this strategy include sample-to-sample variation in chromatographic separation of the positional isomers, uneven matrix and acetonitrile effects due to different retention times and high operating skills required for elaborated chromatography systems.

We developed an easy to implement strategy involving iterative loops of targeted MS1-MS2-MS3 scans. To increase the dynamic range of MS3 acquisition, we performed enhanced zoom scans in the ion trap, targeting an increased window centred on the second isotope of the di-acetylated MS2 fragment ions – but excluding the mono-acetylated precursor ions – and applying WideBand activation to increase the efficiency of fragmentation. Note that the second isotope of the MS2 y12 ion

that is subjected to MS3 (theoretical mass: 1215.7096) is often at least as strong as the first isotope (see Figure S1A second panel zoom in). Although a number of MS3 fragment ions could potentially be analysed, we focused on the b10 ion pairs because this gave most reproducible results. The b10 ions with two acetyl groups represent H4.K8acK12ac (theoretical $m/z = 867.469$) and b10 ions with a single acetyl-group represents H4.K8acK16ac and H4.K12acK16ac (theoretical $m/z = 870.492$). Note that fragmentation of the water loss ion through WideBand activation shifted the observed spectra by -18 Dalton (see Figures S2A, S2C). Supplementary Table S2 details the formulas that are used to calculate the motif abundances for the di-acetylated isoforms.

Theoretically, the b10 MS3 fragment ion with two acetyl groups (theoretical $m/z = 870.492$) may interfere with the y8 MS3 fragment ion harbouring a single acetyl group (theoretical $m/z = 870.516$). To test for this, we recorded MS3 scans for a synthetic peptide (H4.K8acK12ac) for which we only expect signals for the b10 MS3 fragment ion with two acetyl groups (expected m/z at 849) and the y8 MS3 fragment with a single acetyl group (expected m/z at 855). The MS3 signal intensities for the b10 ion are dominant over the y8 ion (40400 vs 1526, $b10/y8 > 0.96$), demonstrating the usability of the b10 ion for quantification. Alternatively, one may use the b8 MS3 classifier pair (theoretical m/z at 739.41 and 742.43) yet we observed less accurate results using synthetic peptides and overall less signal and more varying results with most biological samples.

To test for accuracy, we performed sets of experiments where we mixed synthetic peptides containing di-acetylated H4 motifs in different ratios and measured their abundances via LC-MS. The representative experiment in Figure S3B shows that the measured ratios are close to the expected ones, verifying the MS3-based assay and demonstrating high accuracy.

In summary, we introduce an easy to implement and robust LC-MS workflow which resolves the abundance of PTMs within complex positional isomers. The assay is general applicable and might be useful to pinpoint acetylation sites within clusters of any protein.

1.3 Detection and quantification of H3 and H4 histone acetylation motifs

Histones are basic proteins with clusters of lysines close to the N-terminal ‘tail’ domains. The goal of this study was to monitor the acetylation status of any lysine on the canonical histones H3 and H4 with a particular focus to probe for potential combinatorial motifs. Because we chemically acetylated lysines which are not acetylated in the cell, peptides from endogenously acetylated and non-acetylated lysines co-elute during chromatography. This procedure not only has the advantage of increasing the precision (see above), but also to increase the confidence in the identification of low abundant acetylation sites. Although the MS signals of low abundant acetylation sites often display more imprecise masses than high-signal MS identifiers, precise co-elution (or slight shift towards later retention times) is a good indication for the existence of an acetylation site (see below for examples

including H3.K27me2K36ac, H3.K36me2/3K27ac) whereas the absence of co-elution is a strong indication for the absence of an acetylation site (see H4.K20ac and H3.K56ac below). A slight shift towards later retention times for endogenously acetylated histones compared to chemically d3-acetylated histones is caused by the slightly decreased interaction of the deuterated (d3) moieties with the C18 column (Goodlett et al., 2001).

We robustly detected MS1 masses corresponding to the mono-, di-, tri- and tetra-acetylated isoforms of the H4.G4-R17 peptide from KC cell histone preparations (Figure S2A). MS2 and MS3 peptide sequencing of histones from KC cells enabled the identification and quantification of all positional isomers on this peptide with one exception (H4.K5acK16ac). Although we detected synthetic peptides to H4.K5acK16ac and accurately determined their abundance in mixtures of synthetic di-acetylated H4.G4-R17 peptides (Figure S2B), the low abundance of this motif together with the more complicated formula to derive this motif (see Table S2) did not allow its reliable quantification in KC cell samples. We therefore do not display the abundance values for this motif. Moreover, to keep the other di-acetylated H4.G4-R17 motifs independent of H4.K5acK16ac, we did not artificially correct for omitting H4.K5acK16ac. Therefore, if we sum up the abundances of all fifteen quantified H4.G4-R17 motifs, we only approximate but do not exactly arrive at 1.

Although we detect all histone H4 peptides in their unmodified form, we did not identify any acetylation to lysines K31, K44, K59, K77, K79 and K91 (Figure S2). Acetylation at H4.K20ac has been detected in plants (Zhang et al., 2007) and recently also in human cancer cells (Zheng et al., 2013; Tang et al., 2014). In contrast, we fail to detect H4.K20ac in *Drosophila* cells. Of note, in a few histone preparations from adult flies we detected a MS1 mass at $m/z = 279.1940$ which is very close to the theoretical mass of H4.K20ac ($m/z = 279.1927$). Its very low signal intensity did not allow MS2 analysis. Importantly, because the peaks that may indicate H4.K20ac did not co-elute with (or elute slightly after) the peaks of the unmodified H4.K20-R23 peptide, we conclude that we are currently unable to detect H4.K20ac in *Drosophila* cells.

Similar to H4.K20ac, we also did not detect acetylated H4.K91. H4.K91ac has been found in yeast and mammalian cells and is associated with histone maturation in the cytoplasm (Ye et al., 2005; Yang et al., 2011). Figure S2G shows the existence of a high-signal triple charged MS1 spectrum for the unmodified H4.K79R92 peptide, yet we do not detect any signal for the acetylated isoform at a similar retention time.

Consecutive MS1-MS2 analysis on histone H3 peptides identified acetylation to lysines 9, 14, 18, 23, 27, 36 and 37 in the combinations detailed in Table S4. Figure S3 documents the MS1 and MS2 spectra. Co-elution of H3.K27me2K36ac, H3.K36me2K27ac and H3.K36me3K27ac with the unmodified isoforms H3.K27me2, H3.K36me2 and H3.K36me3 provides high confidence in the identification of these motifs. The b3 and y11 fragment ion pairs derived from targeting the mono-

acetylated H3.K27-R40 peptide revealed that while most acetylation is, as expected, on lysine 27 (0.20% of all H3 motifs corresponding to estimated 12000 molecules/cell), there was also detectable acetylation on lysine 36 (0.058% of H3 motifs or estimated 3500 molecules/cell; for calculations see below). Moreover, we detected minute amounts of acetylation on lysine 37 (0.003% of H3 motifs or estimated 182 molecules/cell, Figure S3R-T). Although the signals of the MS2_K27R40_y4_K36ac_579 identifier reporting on H3.K37ac is above background, the current number of replicates within the RNAi dataset did not allow its high-confidence quantification. We therefore display only 28 out of 29 acetylation motifs for the KAT and KDAC RNAi datasets. Using the motif abundance of H3.K37ac, we estimate that our lowest detection limit is at around 0.003% or estimated 182 molecules per cell for those motifs that require consecutive MS1 and MS2 analysis. Motifs that require solely MS1 analysis likely have lower detection limits (including H4.K20ac, H3.K56ac).

We did not detect the long hydrophobic peptide H3.F84-R116 and thus cannot interrogate acetylation to lysine H3.K115. Furthermore, although we robustly detected all other unmodified H3 peptides, we did not detect acetylation to lysines K4, K56, K64, K79 and K122 (Figure S3).

Acetylation of H3.K56 is high in yeast cells and very low in mammalian cells (see (Drogaris et al., 2012) and references therein). Using antibodies to detect H3.K56ac, Tyler and colleagues reported that deprivation of SIR2 in *Drosophila* S2 cells increased the levels of H3.K56ac (Das et al., 2009). We did not detect H3.K56ac in KC cells treated with control RNAi or RNAi targeting SIR2 (Figure S3J). We also did not detect this modification in S2 cells. Measuring synthetic peptides of unmodified and acetylated K56 yielded strong signals for both isoforms, arguing that our LC-MS approach is capable to detect acetylated K56ac if present on synthetic peptides (Figures S3K, L).

Acetylation of H3.K64 and H3.K122 was recently observed in mammalian cells (Tropberger et al., 2013; Di Cerbo et al., 2014), but similar to H3.K56ac, we detected the unmodified but not acetylated isoforms in histones from KC cells (Figure S3X, Y).

We detected methylation at lysine 20 of H4 and lysines K4, K9, K27, K36 and K79 of H3 (Table S4 and Figure S2, S3). Although occasionally we observed a mass corresponding to a mono-methylated H3.K9-R17 peptide, which elutes 3 minutes after H3.K9me1 and thus might be H3.K14me1, MS2 analysis did not confirm the identity of H3.K14me1. The same applies to MS1 masses, which may indicate mono-methylation at the H3.K18-R26 peptide. MS2 analysis targeting this parent ion did not yield any fragment ion that would indicate the presence of the mono-methylated H3.K18-R26 peptide.

Apart from the methyllysine-containing motifs detailed in Figure 1D, E and Table S4, we also detected low levels of H3.K36me1, H3.K27me1K36me1 and minute amounts of H3.K79me3. Under our conditions, H3.K36me1 elutes very close to the positional isomer of the more dominant H3.K27me1 (Figure S1B) and in some samples H3.K36me1 ($m/z = 528.3226$) co-eluted with

acetylated K27ac from the histone variant H3.3, which lead to an interference with its second isotope at $m/z = 528.3157$. We therefore currently cannot reliably report H3.K36me1 throughout all KAT and KDAC RNAi samples. We also excluded H3.K27me1K36me1 for a high-confidence comparative analysis, because its measured abundance varied substantially even throughout control RNAi samples. Similar to other di- and tri-methyl pairs, H3.K79me3 elutes shortly before H3.K79me2 (Figure S1A, lowest panel in schedule 4). Because of its very low signal intensity and the observed variability across control RNAi samples, we excluded H3.K79me3 for a high-confidence quantitative analysis. We estimated the abundance of H3.K79me3 in KC cells to be below 0.2% (Figure S3W), supporting that, similar to human DOT1L but unlike to yeast Dot1, *D. melanogaster* DOT1 almost exclusively generates H3.K79me1/2.

In this study, we do not report on the methylation states of H4.K20 and H3.K4. Although we robustly identified the more hydrophobic un-methylated and mono-methylated states to each of these peptides (Figure S2D-F, S3A,B), H3.K4me2 showed much lower signal intensities and the me3 states of each peptide are only occasionally detected. Reliable quantification of these small hydrophilic peptides would require additional or alternative peptide purification strategies. To not compromise quantification of the other 45 histone PTM motifs, we excluded the analysis of the H3.T3-R8 and H4.K20-R23 peptides.

Calculation of motif abundance and estimation of cellular copy number

To calculate the relative abundance of a histone motif, its specific peak intensity (area under curve of each MS chromatogram) was divided by the sum of all peak intensities of motifs that contribute to a peptide (See Table S2). For example, the relative abundance of H3.K14ac was calculated by multiplying the fraction of MS2 identifiers that distinguish H3.K14ac from H3.K9ac ($\text{MS2_K9R17_y7_K14ac_728} / (\text{MS2_K9R17_y7_K14ac_728} + \text{MS2_K9R17_y7_K14noAc_731})$) with the fraction of the non-methylated but mono-acetylated MS1 signals relative to all MS1 signals for the H3.K9-R17 peptide ($\text{H3.K9-R17_1Ac} / (\text{H3.K9-R17_noPTM} + \text{H3.K9-R17_1Ac} + \text{H3.K9-R17_2Ac} + \text{H3.K9-R17_K9me1/2/3} \pm \text{K14ac})$). Of note, any other modification that occurs on this motif but was not quantified in this study will cause an underestimation of the abundances for the quantified motifs. This includes, for instances, the phosphorylation sites at H3.S10 and H3.T11 – which would require alternative purification strategies – as well as PTMs that have not yet been identified.

We attempted to determine the exact cell copy number of all histone molecules and specific histone acetylation and methylation motifs by spiking-in highly purified, precisely quantitated synthetic peptides (TQL-spiketide, JPT Berlin) into known quantities of KC cells. However, as described below, we observed already strong variations in the abundance of the Q tag dissociated by trypsin digestion

from different proteotypic peptides, which complicates its faithful accurate quantitative interpretation. Moreover, we observed strong variations of the LC-MS response throughout replicate experiments with cells and spike-in TQL-peptide mixes. Lastly, we observed ion suppression effects when we titrated the number of cells against the concentration of TQL synthetic peptides, or *vice versa*.

Because of the technical problems described above, we estimated the number of histone motifs per cell assuming i) a *D. melanogaster* genome size of 196 mega base pairs (Mb) for females and 170 Mb for males (<http://flybase.org/reports/FBsp000000001.html>), ii) an average nucleosome repeat length of 197 base pairs (Becker and Wu, 1992), iii) a minimal pool of free histones while most histones are organised as nucleosomes (Loyola et al., 2006) and iv) equal number of cells in G1 (n=2) and G2 (n=4) phase. Accordingly, multiplying the relative abundances measured in this study with 5 969 543 and 5 177 664 million histone molecules of a diploid female or male cell, respectively, the estimated abundance for the rare H3.K37ac motif is around 182 molecules per female cell whereas the high abundant H3.K23ac motif is present on approximately 2.8 million histone molecules.

Supplementary Note 2: Introducing an LC-MS response factor to increase the accuracy of histone PTM quantification

Post-translational modifications can have a substantial impact on several steps of the LC-MS workflow – including column interactions during liquid chromatography and ionization and fragmentation within the mass spectrometer – and can therefore severely bias accurate quantitation between different PTMs signatures of the same peptide (Schotta et al., 2008; Farley and Link, 2009; Marx et al., 2013). For the current study, we needed to know accurate relative abundances to reliably interpret the abundance levels in the light of their roles as high abundant structural motifs or rare motifs with signalling character.

As expected, the D3AA method applied in this study reduced the inherent bias in MS signals for motifs that contain a single or multiple acetylations as the only PTM. Measuring defined amounts of different synthetic di-acetylated H4.G4-R17 peptides recovered the expected ratios (see above and Figures S2B, C). Furthermore, we observed very similar LC-MS response factors among positional isomers of H3 peptides that contain lysine acetylation as the only PTM (Table S3).

In contrast, we observed dramatically different MS signals when we measured the same amount of peptides that differed in their methylation status. A similar observation has been made before by Jenuwein and colleagues who observed up to 5 fold differences when comparing the raw MS signals among synthetic peptides to H4.K20me1/2/3 and the unmodified H4.K20-R23 peptide (Schotta et al., 2008). We measured synthetic peptides to methylation motifs that contain methylated K9, K27, K36 and K79. The most dramatic effect was measured between the different methylation states of lysine 9,

where the raw MS signals of the H3.K9me3 peptide was 50-fold less than the raw MS signals for the H3.K9me1 motif (Table S3, column 5).

Some of the differences might be due to different *initial* peptide concentrations, which may vary in a peptide-specific manner due to peptide processing workflows between the quantitation of peptides (at JPT, Berlin) and their digestion with trypsin (in our laboratory). For example, the exact quantitation of the TQL-synthetic peptide by fluorescence might be influenced by the physicochemical properties of the peptide sequence including their modification pattern. Furthermore, peptides with different physicochemical properties, such as the ones with different methylation states, might be differentially affected by peptide loss during the early pre-LC workflows, including insufficient resuspension and different interactions with the tube walls.

To test for this ‘pre-LC error’, we measured the MS signals of the Q tag, which is cleaved off the proteotypic peptide by trypsin digestion. Indeed, we observed substantial differences in the Q tag signals between experiments with different TQL-synthetic peptides (Figure 1C, middle panel). Normalising the MS signals of the proteotypic peptides by the MS signals of the Q tag not only reduced the variation between replicate measurements by two-fold (compare the SEM between columns 4 and 6 in Table S3) but also yielded more similar LC-MS response correction factors among motifs that differ only in their site of acetylation. For example, while the LC-MS response correction factors without applying the Q tag normalisation differed markedly between H3.K9ac (3.92), H3.K14ac (1.62) and H3.K9acK14ac (2.26), applying the normalisation with the Q tag signals converged the LC-MS response correction factors (H3.K9ac: 1.46, H3.K14ac: 1.46 and H3.K9acK14ac: 1.58). In addition, the Q tag normalisation also reduced the overall differences between the highest and lowest LC-MS response correction factor (e.g. H3.K9me3 before Q tag normalisation: 50.49, after Q tag normalisation: 34.17).

Theoretically, the accuracy of the LC-MS response factors for motifs that elute during different retention times can be increased by measuring them in the background of the endogenous histone sample. However, because we observed a dose-dependent interference between the synthetic spike-in peptides with their co-eluting endogenous counterpart, we did not proceed with this strategy, which is also the reason why we refrained from using the spike in peptides for absolute quantitation.

Application of the LC-MS response correction factor had a significant impact on estimating the cellular copy number of histone acetylation and methylation motifs (Table S4). For example, without applying the LC-MS response correction factor, we estimated that H3.K27ac is found on 0.7% of H3 molecules (estimated 41,000 molecules/cell). This number is unexpectedly high, considering that this dynamic mark is preferentially detected at active enhancers (> 10,000) (Kharchenko et al., 2011; Yanez-Cuna et al., 2014). If we correct for the LC-MS bias, we arrive at a 3.5 times lower number for H3.K27ac (0.2% or estimated 12,000 molecules/cell). By contrast, we underestimated the abundance

of the repressive H3.K9me3 mark by almost ten-fold (4.2% vs. 39%). Combining the numbers for the cellular abundance, genome-wide occupancy and half-life allows deriving testable quantitative models on genome organisation and function.

Application of the LC-MS response correction factor did not change the conclusions of most KAT and KDAC RNAi experiments, with a few notable exceptions (Figure S6 and Table S5). While H3.K14ac appeared to be elevated in many KAT depletion experiments without applying the LC-MS correction factor, this trend was weaker when correcting for the LC-MS bias. Another example is the depletion of HBO1, that appeared to cause a moderate reduction of H3.K9me2K14ac (85% of control), which was not observable anymore after applying the correction factor (98% of control).

Supplementary Note 3: Experimental design and statistical analysis

During the course of this study, we continuously developed the LC-MS protocol while expanding the scope of KAT and KDACs and histone motifs to be analysed. Therefore, the current manuscript summarises datasets that were generated using different LC-MS workflows and different KC cell batches grown and harvested during different months and years (time span from 2012 to 2014). As a consequence, we observed technical and biological variations, which we addressed by normalising KAT/KDAC RNAi samples within each ‘batch’ to their corresponding control RNAi samples (Tables S4, S5 and see below).

We aimed to ablate each KAT and KDAC with two distinct non-overlapping RNAi constructs (Table S7) in at least two independent biological experiments for each RNAi construct. For most RNAi experiments, we used two dsRNA sequences from the *Drosophila* RNAi Screening Center (DSRC, Harvard). Those constructs have an average size of 500 base pairs and are widely used in genome-wide RNAi screens. For some genes, we designed new RNAi constructs using the NEXT-RNAi platform (www.nextrna.org) because there was only a single RNAi construct at DSRC (e.g. MGEA5) or we expected an insufficient knockdown or elevated off-target reactivity when using short DSRC constructs (e.g. NAA10 < 250 bp). Because of their short mRNA size or unfavourable sequence composition, we could only target NAT9, NAA40, CG12560 and SIRT6 with a single dsRNA construct.

For most RNAi experiments, we observed an overall high similarity of the responses that are caused by different dsRNA constructs targeting the same KAT or target gene (Table S5). However, this was not the case for each one dsRNA-CBP and dsRNA-TIP60 construct, which were therefore not incorporated in the analysis (see note in Table S5).

While many KAT RNAi samples were acquired during three different ‘replicate batch groups’ (see Table S5), some KAT RNAi samples were acquired in fewer batches (e.g. MGEA5: 3 biological replicates in the ‘C’ batch; NAT10: 2 biological replicates each in the ‘B’ and ‘C’ batch). In some

cases, we conducted several independent RNAi experiments with the same batch of cells yet on different days or weeks. For example, we conducted two independent RNAi experiments each with the single NAA40_1 RNAi construct during times of the ‘A’ batch (NAA40_1_1_A and NAA40_1_2_A) and the ‘B’ batch (NAA40_1_1_B and NAA40_1_2_B).

Normalisation to control RNAi samples within each batch was necessary because, while most histone motifs did not show major differences in their abundance across the different batches, some motifs showed considerable variation. For example, while the abundance of H4.K5ac was similar across the batches ‘A’ to ‘E’ (3.38%, 3.42%, 3.75%, 3.15% and 3.38%), the abundance values scattered stronger for H4.K12ac (6.95%, 8.20%, 7.66%, 5.98% and 8.03%). Because our interest was primarily in the relative change of a motif comparing a KAT/KDAC RNAi with its control (GST RNAi, GFP RNAi, EGFP RNAi or mock), this normalisation step reduced biological variation across *paired* groups of KAT/KDAC and control RNAi (= samples between ‘batches’) and thereby increased statistical power to detect minor differences. Although we note that some motif abundances systematically varied more than others, we currently suspect not only technological reasons (such as noise within LC-MS signals due to different quantitation requirements, see Table S2) but also biological explanations, such as the time we kept cells in culture or environmental conditions. For example, we observed varying levels of H4.K12ac and H3.K18acK23ac in cells, which were either kept for very short or extended periods of time in culture (preliminary observation, samples not used during this study). Similarly, we suspect altered serum conditions to be responsible for some batch-to-batch variations.

We excluded MS values for histone motifs that we either did not measure in a particular sample or in case the MS values were not reliable. For example, we did not measure histone H3 motifs of samples within batch ‘A’. Common sources for non-reliable MS values were low MS signals due to technical shortcomings at the LC or MS, interference with other ions (sometimes caused by problems with the LC gradient) or ion suppression effects. In case we identified a problem with a single MS identifier (e.g. the early eluting MS1_K9R17_K9me3 identifier, which reports on the H3.K9me3 motif), we excluded all motifs of this peptide (H3.K9-R17), even if individual motifs, which elute at later retention times, appeared fine (H3.K9me1 and acetylated K9 and K14). However, in many cases, other peptides were not affected (e.g. H3.R18-K26) and thus remained in the analysis. In these special cases where we could not reliably identify peptides, we did not calculate summarised values (e.g. global acetylation score).

Supplementary Note 4: Mining the database to derive testable hypotheses and shed light into controversial enzyme-substrate relationships

HAT1 is the well-known cytoplasmic HAT that renders histone H4 competent for nucleosome assembly by acetylation at lysines 5 and 12. Up to now, current methodology did not allow measuring

the di-acetylated motif *per se* (see Suppl. Note 1.2). Our analysis shows that the main effect of HAT1 knockdown is a reduction of the H4.K5acK12ac mark (Figure 2A). Interestingly, it also reveals a reduction of K5ac, but not of K12ac. Together with *in vitro* data on recombinant HAT1 that demonstrates efficient acetylation of unmodified and pre-acetylated templates on either positions 5 or 8 yet not on other sites or combinatorial motifs (K12ac, K5acK12ac, K8acK16ac; Makowski et al. 2001), this leads us to hypothesise that HAT1 first acetylates H4 at K5, releases a fraction of H4.K5ac to the cellular histone pool and di-acetylates another fraction to produce H4.K5acK12ac. Notably, we detect a minor decrease of the tri-acetylated H4.K5acK8acK12ac upon HAT1 depletion, which suggests that HAT1 further acetylates K8 on a small fraction of its di-acetylated product. This is in line with previous reports indicating that cytosolic H4 bound to chaperones is not only acetylated at lysines 5 and 12, but that a minor acetylation of K8 can be detected as well (Verreault et al., 1996).

The physiological substrates of HBO1 are controversially discussed. Whereas some groups report HBO1 as the major H4 acetyltransferase with specificity for K5, K8 and K12, others found that inactivation of the *hbo1* gene in mice did not cause a decline of H4 acetylation, yet led to dramatic loss of H3.K14ac (Miotto and Struhl, 2010; Kueh et al., 2011; Lalonde et al., 2013). Indeed, HBO1 ablation in *Drosophila* affects H4.K12ac levels only very modestly, if the values of all motifs including this mark are combined (as if one used an antibody that does not differentiate H4.K12ac-containing motifs) (Figure 2B). However, monitoring individual motifs reveals that HBO1-lacking cells show a moderate reduction of mono-acetylated H4.K12, but a severe loss of the di-acetylated H4.K8acK12ac motif (Figures 2A, 3A, 3B). One of the possible scenarios to explain this finding is that HBO1 initially acetylates H4.K12 and then further acetylates K8 on a fraction of its first acetylation product.

Supplementary Note 5: Homeostatic control of acetylation levels upon deprivation of KATs

To determine whether deprivation of KATs reduces the global acetylation levels of histones, we calculated a ‘global H3 + H4 acetylation score’ and compared it between cells lacking a specific KAT and cells treated with control RNAi constructs (Figure 5A and Table S5, column 59). The ‘global H3 + H4 acetylation score’ is calculated by summing up all fractional acetylation events observed on the four histone peptides that are subject to acetylation (H4.G4-R17, H3.K9-R17, H3.K18-R26 and H3.K27-R40). For example, depletion of CBP in the CBP_2_3_C experiment (targeting CBP with the CBP_2 RNAi construct, using the third biological replicate experiment in the batch ‘C’) displayed a relative ‘global H3 + H4 acetylation score’ of 96% (Table S5, row 74, column 59). In other words, the cumulative acetylation on the four acetylated peptides (H4.G4-R17, H3.K9-R17, H3.K18-R26 and H3.K27-R40) was reduced by only 4% relative to control RNAi. This value is derived by dividing the ‘global H3 + H4 acetylation score’ for CBP_2_3_C of 100.57 (see row 74, column 117) by the mean

of the ‘global H3 + H4 acetylation score’ for the control RNAi samples from the ‘C’ batch, which has a value of 105.02 (Table S5, row 266, column 117). The ‘global H3+H4 acetylation score’ (column 117) is the sum of the ‘global acetylation score for H4’ (column 115) and the ‘global acetylation score for H3’ (column 116). The ‘global acetylation score for H4’ represents the fraction of histone H4 peptides that is acetylated, which is 27.14% for the CBP_2_3_C sample. The ‘global acetylation score for H3’ is the sum of all fractional acetylation events on the three acetylated H3 peptides (19.78% for H3.K9-R17, 53.59% for H3.K18-R26, and 0.06% for H3.K27-R40, total of 73.43%).

To calculate the % total acetylation change that is contributed by individual motifs (Figure 5B), we scaled the relative change of each KAT RNAi/ctr RNAi experiment (filtered for significant losses, using unpaired two-sided t test on log2 ratios) to the relative abundance of each motif in KC cells. This was done to circumvent the batch effects between different ‘replicate batch groups’ (see Suppl. Note 3).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Inventory for KATs and KDACs in *Drosophila*

A list of putative KATs and KDACs was generated using public databases (FlyBase, ENSEMBL, H1stome (Histone Info Database – ACTREC), HomoloGene), primary research articles, reviews and cross-homology search using BLAST. Gene expression data was obtained from FlyBase (FlyAtlas, modENCODE tissue expression data, modENCODE temporal expression data, modENCODE cell line expression data).

Fly stocks, crosses and handling

Flies were raised on standard cornmeal/yeast medium at 18°C - 22°C. OregonR yw flies were used and considered wildtype. The *mof2* allele is described in (Gu et al., 1998) and the MOF(+) rescue transgene is published in (Prestel et al., 2010). Because we did not observe phenotypic effects for female *mof2/mof2* mutants under our culturing conditions, we maintained them as stable stocks and used directly 20 - 50 females either 1 day or 2 - 4 days old for histone PTM analysis. Hemizygous male *mof2* mutants are lethal at the larval stages, as described earlier (Gu et al., 1998), and therefore cannot be analysed at the adult stages of development. MOF(+)/CyO-GFP flies were generated and those males were crossed to homozygous female virgins *mof2/mof2*; +/+ to obtain *mof2/y*; +/CyO-GFP hemizygous male mutants. GFP fluorescence was used to sort male third-instar larval *mof2* mutants.

Cell line cultivation and RNA interference (RNAi)

Cultivation of KC and S2 cells and RNAi were carried out essentially as described before (Feller et al., 2012). Briefly, cells were grown in T75 or T175 flasks, washed in PBS and 1.5 million cells were transferred to each well of a 6-well plate. Cells were treated with 10 µg dsRNA for 60 min in 1 ml serum-free medium. After addition of 2 ml serum-containing medium (10% FCS), cells were incubated for 5.5 days before harvest. Primer sequences used for RNAi are described in Table S7. We used S2 cells of the ‘L2-4’ subtype, which were obtained from Dr. Patrick Heun (MPI Freiburg, Germany).

Human HeLa Kyoto cells were grown in DMEM medium containing 10% FCS and 1% penicillin/streptomycin at 37°C and 5% CO₂. Cells were transfected with siRNAs using oligofectamine (Invitrogen) according to the manufacturer’s instructions. Growth medium was exchanged every day and cells were harvested for histone PTM analysis after 2.5 days. Primer

sequences used for RNAi are described in Table S7 and have been published in (Taipale et al., 2005) and (Smith et al., 2005).

RT-qPCR experiments

Total RNA was prepared from KC cells using Trizol followed by the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. 600 ng of total RNA was reverse transcribed using random primers and the SuperScript III kit (Life Technologies). Real-time PCR was performed with a Roche 480 light cycler using the second derivative method for quantification. Primer sequences are described in Supplementary Table S7.

Immuno blotting

Quantification of immunoblots was performed using the LI-COR system (Biosciences). The lamin antibody was obtained from Harald Saumweber (Berlin) and the MOF antibody was described before (Prestel et al., 2010)

Extraction of histones from *Drosophila* tissue and cell lines and peptide processing

Third-instar male larvae were extensively washed in PBS, collected in microcentrifuge tubes, flash frozen in liquid nitrogen and stored at -80°C. Adult female flies were processed in the same manner but the washing step was omitted. Frozen tissue was homogenised in A1 buffer (15mM HEPES pH 7.6, 60 mM KCl, 15 mM NaCl, 4 mM MgCl₂, 0.5% Triton, 0.5mM DTT, protease inhibitor (Roche)), centrifuged and the pellet was resuspended in 0.2 M sulphuric acid. Acid extraction and all downstream steps were essentially done as with tissue culture cells, as follows.

KC and S2 cells were harvested and collected by centrifugation, washed once with PBS and pellets were flash frozen in liquid nitrogen and stored at -80°C. Frozen cell pellets were suspended in 0.2 M sulphuric acid and histones and other acid-soluble proteins were extracted overnight and precipitated with 26% TCA. After two to five washes with cold acetone, the histone pellet was resuspended in SDS-PAGE loading buffer. If necessary, pH was adjusted with 1 µl of 1 M Tris pH 8.0. Histones were separated by SDS-PAGE on 15% polyacrylamide gels, stained with Colloidal Blue Staining Kit (Invitrogen) and bands corresponding to histones H3 and H4 were excised from the gels.

Gel pieces containing histones were washed twice with water, twice with 100 mM ammonium bicarbonate (ambic) and incubated in 50 mM ambic/50% acetonitrile for 3 times 10 min while shaking at 37°C for destaining. Gel pieces were successively dehydrated by incubating once with 100 mM ambic, once with 20 mM ambic and 3 times with acetonitrile. Histones were chemically acetylated

with d6-deuterated acetic anhydride (99% D, Sigma) for 45 min at 37°C (pH was monitored to be pH 7-8). After acetylation, histones were washed 4 times with 100 mM ambic and thrice with acetonitrile followed by overnight trypsin (Promega) digestion at 37°C. Tryptic peptides were extracted twice with 70% acetonitrile/0.25% TFA and twice with acetonitrile, vacuum concentrated and resuspended in 0.1% TFA. Histone peptides were desalted using C18-StageTips (Rappsilber et al., 2003), eluted in 80% acetonitrile/0.25% TFA, vacuum concentrated, reconstituted in 0.1% TFA and stored at 4°C (short) or -20°C (long).

Processing of TQL spiketides and calculation of LC-MS response correction factor (LCF)

Purified, isotopically labelled and quantified SpikeTides TQL were purchased from JPT Peptide Technologies GmbH (Berlin). 1 nmol of peptides was resuspended in 80% 100 mM ambic, 18% acetonitrile and 2% DMSO, followed by two rounds of 1 min vortexing, 3 min sonication and 15 min shaking at 25°C. Chemical acetylation and trypsinisation was carried out as described above. Each TQL proteotypic peptide was measured in replicates in different concentrations (1, 5 and 15 pmol) using the same LC parameters but acquiring MS1 scans only.

The LC-MS response correction factor (LCF) was calculated by dividing the raw MS signals for the proteotypic peptide per pmol (column 5) by the raw MS signals for the Q tag which was followed by scaling to the motif that has the highest signal among all motifs for a given peptide. We experienced technical problems for individual motifs from the H3.K27-R40 and H3.E73-R83 synthetic peptides and therefore can only report LCFs with moderate confidence for those motifs.

LC-MS workflow and peak integration

Tryptic peptides were injected in two different HPLC systems from Dionex depending on the batch: Samples from replicate batch 'A' were separated on an Ultimate and replicate batches 'B' to 'D' were separated on an Ultimate 3000 RSLCnano. Peptides were separated with a gradient from 5-60% acetonitrile in 0.1% formic acid over 40 min at 300 nl/min on a C18 analytical column (75 µm i.d. ×15 cm, packed in-house with Reprosil Pur C18 AQ 2.4 µm; Doctor Maisch). The effluent from the HPLC was directly electrosprayed into an LTQ-Orbitrap Classic mass spectrometer (Thermo Fisher Scientific). MS was operated in a targeted setup (see Suppl. Note 1). Typical MS conditions were spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 200°C; MS1 resolution of 7500 (at 400 m/z); normalised collision-induced dissociation energy 35% (MS2) and 45% (MS3); activation q = 0.25; and activation time = 30 ms.

LC-MS data was quantified using the Xcalibur software package (Thermo Fisher Scientific, version 2.2) based on the area of the peak from the extracted ion chromatogram. A processing method was

built using the Xcalibur Processing Setup including the mass-to-charge range and the expected retention time (see Table S1). Further parameters were: Peak detection: Genesis; trace: mass range; smoothing points: 3, S/N threshold: 0.5. The identities of the MS quantifiers were verified by MS2 spectra of the natural occurring forms and their corresponding synthetic peptides (JPT Berlin). All chromatograms were manually verified in Quan Browser and re-integrated where necessary. Typically, this involves baseline adjustment and adjustments to integrate the identical retention time across the d0/d3 pair (i. e. endogenous vs. chemical acetylation). Peak integration was optimised to yield the maximal co-eluting chromatogram of the d3/d0 pair. The spectra from suspicious chromatograms (e.g. if peaks with similar m/z were eluting close-by) were manually inspected, and, if necessary, m/z integration range was adjusted (if possible) or MS quantifier were excluded (see Suppl. Note 3).

Data analysis and statistical analysis

After peak integration, the data was exported into an Excel spread sheet and data summarisation and statistical analysis was performed in Excel and R. The relative abundance was calculated according to the formulas provided in Table S2. Each sample was normalised to control RNAi samples within the same biological replicate batch (for details and justification: see Suppl. Note 3) and ratios were log2 scaled. We applied a two-sided unpaired t test over the log2 ratios of relative changes from all target RNAi samples (e.g. HAT1) for a specific motif (e.g. H4.K5acK12ac) relative to all control RNAi samples within the same replicate batch. Because we considered the purpose of this study in the discovery of changed histone motif abundances upon KAT and KDAC RNAi and we expected only moderate differences, we did not apply a conservative filter by adjusting the p--values by their FDRs. Nevertheless, for convenience of the individual reader, Table S5 provides all MS raw signals, relative and total abundance values to all RNAi samples as well as summarised tables, which report on the mean changes of every KAT/KDAC RNAi (unfiltered) and raw and adjusted p values (two-sided unpaired t test, p values were adjusted per KAT/KDAC RNAi across all motifs according to Benjamini and Hochberg; p-value adjustment was performed in R). This table might be used to discover a potential change of a histone motif upon an individual KAT knockdown and to evaluate its significance considering the number of replicates and technical and biological variation.

Table S1: List MS quantifier

MS identifier	theoretical m/z	charge state	sequence	typical retention time [min]	XIC range (min)	XIC range (max)
MS1_T3R8_NoPTM	375.224	2	TKQTAR	17	375.07	375.37
MS1_T3R8_K4me1	382.233	2	TKme1QTAR	18.2	382.08	382.38
MS1_T3R8_K4me2	366.722	2	TKme2QTAR	16.8	366.57	366.87
MS1_T3R8_K4me3	373.730	2	TKme3QTAR	16.4	373.58	373.88
MS1_K9R17_noPTM	496.292	2	KSTGGKAPR	20.7	496.14	496.44
MS1_K9R17_1Ac	494.783	2	K(ac?)STGGK(ac?)APR	20.7	494.63	494.93
MS1_K9R17_2Ac	493.274	2	KacSTGGKacAPR	20.7	493.12	493.42
MS1_K9R17_K9me1	503.301	2	Kme1STGGKAPR	25.8	503.15	503.45
MS1_K9R17_K9me1_K14ac	501.792	2	Kme1STGGKacAPR	25.8	501.64	501.94
MS1_K9R17_K9me2	487.793	2	Kme2STGGKAPR	17.65	487.64	487.94
MS1_K9R17_K9me2_K14ac	486.284	2	Kme2STGGKacAPR	17.65	486.13	486.43
MS1_K9R17_K9me3	494.801	2	Kme3STGGKAPR	17.55	494.65	494.95
MS1_K9R17_K9me3_K14ac	493.291	2	Kme3STGGKacAPR	17.55	493.14	493.44
MS2_K9R17_y7_K14ac_728	728.41	2	TGGK14acAPR	20.7	727.91	728.91
MS2_K9R17_y7_K14noAc_731	731.43	2	TGGK14APR	20.7	730.93	731.93
MS1_K18R26_noPTM	538.837	2	KQLATKAAR	38	538.69	538.99
MS1_K18R26_1Ac	537.328	2	K(ac?)QLATK(ac?)AAR	38	537.18	537.48
MS1_K18R26_2Ac	535.819	2	KacQLATKacAAR	38	535.67	535.97
MS2_K18R26_b2_K18ac_299	299.17	2	K18acQ	38	298.67	299.67
MS2_K18R26_b2_K18noAc_302	302.19	2	K18Q	38	301.69	302.69
MS1_K27R40_noPTM	523.651	2	KSAPSTGGVKKPHR	40	523.55	523.75
MS1_K27R40_1Ac	522.643	2	KSAPSTGGVKKPHR_1Ac	40	522.54	522.74

MS1_K27R40_K27me1	528.323	2	Kme1SAPSTGGVKKPHR	41.6	528.17	528.47
MS1_K27R40_K27me2	517.983	2	Kme2SAPSTGGVKKPHR	35.1	517.83	518.13
MS1_K27R40_K27me2_K36ac	516.975	2	Kme2SAPSTGGVKacKPHR	35.1	516.83	517.13
MS1_K27R40_K27me3	522.655	2	Kme3SAPSTGGVKKPHR	34.8	522.51	522.81
MS1_K27R40_K36me2	517.983	2	KSAPSTGGVKme2KPHR	37.4	517.83	518.13
MS1_K27R40_K36me2_K27ac	516.975	2	KacSAPSTGGVKme2KPHR	37.4	516.83	517.13
MS1_K27R40_K36me3	522.655	2	KSAPSTGGVKme3KPHR	37.4	522.51	522.81
MS1_K27R40_K36me3_K27ac	521.647	2	KacSAPSTGGVKme3KPHR	37.4	521.50	521.80
MS1_K27R40_K27me2K36me1	522.655	2	Kme2SAPSTGGVKme1KPHR	36.9	522.51	522.81
MS2_K27R40_b3_K27ac_329	329.18	2	K27acSA	40	328.68	329.68
MS2_K27R40_b3_K27NoAc_332	332.21	2	K27SA	40	331.71	332.71
MS2_K27R40_y4_K37NoAc_582	582.36	2	K37PHR	40	578.83	579.83
MS2_K27R40_y4_K37ac_579	579.33	2	K37acPHR	40	581.86	582.86
MS2_K27R40_y5_K27NoAc_752	752.47	2	K36(ac?)K37(ac?)PHR	40	751.97	752.97
MS2_K27R40_y5_K27ac_755	755.49	2	K36K37PHR	40	754.99	755.99
MS1_E73R83_noPTM	690.866	2	EIAQDFKTDLR	53.2	690.72	691.02
MS1_E73R83_K79me1	697.874	2	EIAQDFKme1TDLR	55.4	697.72	698.02
MS1_E73R83_K79me2	682.365	2	EIAQDFKme2TDLR	47.5	682.22	682.52
MS1_E73R83_K79me3	689.373	2	EIAQDFKme3TDLR	47.3	689.22	689.52
MS1_K20R23_noPTM	280.705	2	KVLR	23.2	280.56	280.86
MS1_K20R23_K20me1	287.712	2	Kme1VLR	31.50	287.56	287.86
MS1_K20R23_K20me2	272.203	2	Kme2VLR	18	272.05	272.35
MS1_G4R17_noPTM	725.957	2	GKGGKGLGKGGAKR	41	725.81	726.11
MS1_G4R17_1Ac	724.446	2	GKGGKGLGKGGAKR_1Ac	41	724.30	724.60
MS1_G4R17_2Ac	722.934	2	GKGGKGLGKGGAKR_2Ac	41	722.78	723.08
MS1_G4R17_3Ac	721.422	2	GKGGKGLGKGGAKR_3Ac	41	721.27	721.57

MS1_G4R17_4Ac	719.911	2	GKacGGKacGLGKacGGAKacR	41	719.76	720.06
MS2_G4R17_1Ac_y5_K16ac_530	530.31	2	GGAK16acR	41	529.81	530.81
MS2_G4R17_1Ac_y5_K16NoAc_533	533.33	2	GGAK16R	41	532.83	533.83
MS2_G4R17_1Ac_y7_K16K12.1Ac_760	760.46	2	GK12(ac?)GGAK16(ac?)R	41	759.96	760.96
MS2_G4R17_1Ac_y7_K16K12.0Ac_763	763.48	2	GK12GGAK16R	41	762.98	763.98
MS2_G4R17_1Ac_y12_K16K12K8.1Ac_1217	1217.73	2	GGK8GLGK12GGAK16R_1Ac	41	1217.23	1219.23
MS2_G4R17_1Ac_y12_K16K12K8.0Ac_1220	1220.76	2	GGK8GLGK12GGAK16R	41	1220.26	1222.26
MS2_G4R17_2Ac_y5_K16.1Ac_530	530.31	2	GGAK16acR	41	529.81	530.81
MS2_G4R17_2Ac_y5_K16.0Ac_533	533.33	2	GGAK16R	41	532.83	533.83
MS2_G4R17_2Ac_y7_K16K12.2Ac_757	757.43	2	GK12acGGAK16acR	41	756.93	757.93
MS2_G4R17_2Ac_y7_K16K12.1Ac_760	760.46	2	GK12(ac?)GGAK16(ac?)R	41	759.96	760.96
MS2_G4R17_2Ac_y7_K16K12.0Ac_763	763.48	2	GK12GGAK16R	41	762.98	763.98
MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214	1214.71	2	GGK8GLGK12GGAK16R_2Ac	41	1214.21	1216.21
MS2_G4R17_2Ac_y12_K16K12K8.1Ac_1217	1217.73	2	GGK8GLGK12GGAK16R_1Ac	41	1217.23	1219.23
MS2_G4R17_3Ac_y5_K16ac_530	530.31	2	GGAK16acR	41	529.81	530.81
MS2_G4R17_3Ac_y5_K16NoAc_533	533.33	2	GGAK16R	41	532.83	533.83
MS2_G4R17_3Ac_y7_K16K12.2Ac_757	757.43	2	GK12acGGAK16acR	41	756.93	757.93
MS2_G4R17_3Ac_y7_K16K12.1Ac_760	760.46	2	GK12(ac?)GGAK16(ac?)R	41	759.96	760.96
MS2_G4R17_3Ac_y12_K16K12K8.3Ac_1211	1211.69	2	GGK8acGLGK12acGGAK16acR	41	1211.19	1213.19
MS2_G4R17_3Ac_y12_K16K12K8.2Ac_1214	1214.71	2	GGK8GLGK12GGAK16R_2Ac	41	1214.21	1216.21
MS3_G6R17_2Ac_b10_2Ac_867	867.47	2	GGK8acGLGK12acGGA	41	848.47	849.97
MS3_G6R17_2Ac_b10_1Ac_870	870.49	2	GGK8(ac?)GLGK12(ac?)GGA	41	851.99	852.99

Column Headers

MS identifier: name of the MS identifier, containing MS scan type (MS1, MS2 or MS3); target peptide (first and last amino acid); PTM type and position (e.g. MS1_K9R17_K9me1: MS1 is sufficient to determine mono-methylation on lysine 9) and for MS2/MS3: y/b fragment ion assignment and m/z (e.g. MS2_K9R17_y7_K14ac_728: MS2 on K9R17 peptide that reports on y7 fragment ion (which has m/z of 728) is necessary to determine acetylation on lysine 14)

theoretical m/z: theoretical mono-isotopic mass /charge

charge state: charge state of MS identifier

sequence: target peptide sequence

retention time [min]: typical retention time in minutes

XIC range [min/max]: mass/charge range used for chromatographic peak integration (minimal and maximal value). Note that for individual MS quantifier, the XIC range was adjusted to avoid integration of interfering peaks (e.g. MS1_K27R40_noPTM/1Ac: interfering peak at m/z = 523.79 at similar retention time)

Abbreviations

Kac	- lysine acetylation
Kme1	- lysine mono-methylation
Kme2	- lysine di-methylation
Kme3	- lysine tri-methylation
noPTM	- peptide without PTM, hence all lysines are chemically (d3)-acetylated

Table S2: Calculations to derive motif abundance from MS quantifier

Motifs at histone H4

$$\text{H4.G4R17_noPTM} = \frac{\text{MS1_G4R17_noPTM}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})}$$

$$\text{H4.K5ac} = \frac{\text{MS1_G4R17_1Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} * \left(1 - \frac{\text{MS2_G4R17_1Ac_y12_K16K12K8.1Ac_1217}}{\text{sum}(\text{MS2_G4R17_1Ac_y12_K16K12K8.1Ac_1217}, \text{MS2_G4R17_1Ac_y12_K16K12K8.0Ac_1220})} \right)$$

$$\text{H4.K8ac} = \frac{\text{MS1_G4R17_1Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} *$$

$$\left(\frac{\text{MS2_G4R17_1Ac_y12_K16K12K8.1Ac_1217}}{\text{sum}(\text{MS2_G4R17_1Ac_y12_K16K12K8.1Ac_1217}, \text{MS2_G4R17_1Ac_y12_K16K12K8.0Ac_1220})} - \frac{\text{MS2_G4R17_1Ac_y7_K16K12.1Ac_760}}{\text{sum}(\text{MS2_G4R17_1Ac_y7_K16K12.1Ac_760}, \text{MS2_G4R17_1Ac_y7_K16K12.0Ac_763})} \right)$$

$$\text{H4.K12ac} = \frac{\text{MS1_G4R17_1Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} *$$

$$\left(\frac{\text{MS2_G4R17_1Ac_y7_K16K12.1Ac_760}}{\text{sum}(\text{MS2_G4R17_1Ac_y7_K16K12.1Ac_760}, \text{MS2_G4R17_1Ac_y7_K16K12.0Ac_763})} - \frac{\text{MS2_G4R17_1Ac_y5_K16ac_530}}{\text{sum}(\text{MS2_G4R17_1Ac_y5_K16ac_530}, \text{MS2_G4R17_1Ac_y5_K16NoAc_533})} \right)$$

$$\text{H4.K16ac} = \frac{\text{MS1_G4R17_1Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} * \frac{\text{MS2_G4R17_1Ac_y5_K16ac_530}}{\text{sum}(\text{MS2_G4R17_1Ac_y5_K16ac_530}, \text{MS2_G4R17_1Ac_y5_K16NoAc_533})}$$

$$\text{H4.K5acK8ac} = \frac{\text{MS1_G4R17_2Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} * \frac{\text{MS2_G4R17_2Ac_y7_K16K12.0Ac_763}}{\text{sum}(\text{MS2_G4R17_2Ac_y7_K16K12.0Ac_763}, \text{MS2_G4R17_2Ac_y7_K16K12.1Ac_760}, \text{MS2_G4R17_2Ac_y7_K16K12.2Ac_757})}$$

$$\begin{aligned}
\text{H4.K5acK12ac} = & \frac{\text{MS1_G4R17_2Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} * \\
& \left(\left(1 - \frac{\text{MS3_G6R17_2Ac_b10_1Ac_867}}{\text{sum}(\text{MS3_G6R17_2Ac_b10_1Ac_867}, \text{MS3_G6R17_2Ac_b10_0Ac_870})} \right) * \frac{\text{MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214}}{\text{sum}(\text{MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214}, \text{MS2_G4R17_2Ac_y12_K16K12K8.1Ac_1217})} \right) - \\
& \frac{\text{MS2_G4R17_2Ac_y5_K16ac_530}}{\text{sum}(\text{MS2_G4R17_2Ac_y5_K16ac_530}, \text{MS2_G4R17_2Ac_y5_K16NoAc_533})} - \frac{\text{MS2_G4R17_2Ac_y7_K16K12.0Ac_763}}{\text{sum}(\text{MS2_G4R17_2Ac_y7_K16K12.0Ac_763}, \text{MS2_G4R17_2Ac_y7_K16K12.1Ac_760}, \text{MS2_G4R17_2Ac_y7_K16K12.2Ac_757})} \Big)
\end{aligned}$$

$$\begin{aligned}
\text{H4.K5acK16ac} = & \frac{\text{MS1_G4R17_2Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} * \left(\frac{\text{MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214}}{\text{sum}(\text{MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214}, \text{MS2_G4R17_2Ac_y12_K16K12K8.1Ac_1217})} * \right. \\
& \left. \left(\frac{\text{MS3_G6R17_2Ac_b10_1Ac_867}}{\text{sum}(\text{MS3_G6R17_2Ac_b10_1Ac_867}, \text{MS3_G6R17_2Ac_b10_0Ac_870})} - 1 \right) + \frac{\text{MS2_G4R17_2Ac_y5_K16ac_530}}{\text{sum}(\text{MS2_G4R17_2Ac_y5_K16ac_530}, \text{MS2_G4R17_2Ac_y5_K16NoAc_533})} \right) \Big)
\end{aligned}$$

$$\begin{aligned}
\text{H4.K8acK12ac} = & \frac{\text{MS1_G4R17_2Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} * \\
& \frac{\text{MS3_G6R17_2Ac_b10_1Ac_867}}{\text{sum}(\text{MS3_G6R17_2Ac_b10_1Ac_867}, \text{MS3_G6R17_2Ac_b10_0Ac_870})} * \frac{\text{MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214}}{\text{sum}(\text{MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214}, \text{MS2_G4R17_2Ac_y12_K16K12K8.1Ac_1217})} \Big)
\end{aligned}$$

$$\begin{aligned}
\text{H4.K8acK16ac} = & \frac{\text{MS1_G4R17_2Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} * \left(\frac{\text{MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214}}{\text{sum}(\text{MS2_G4R17_2Ac_y12_K16K12K8.2Ac_1214}, \text{MS2_G4R17_2Ac_y12_K16K12K8.1Ac_1217})} * \right. \\
& \left. \left(1 - \frac{\text{MS3_G6R17_2Ac_b10_1Ac_867}}{\text{sum}(\text{MS3_G6R17_2Ac_b10_1Ac_867}, \text{MS3_G6R17_2Ac_b10_0Ac_870})} \right) - \frac{\text{MS2_G4R17_2Ac_y7_K16K12.2Ac_757}}{\text{sum}(\text{MS2_G4R17_2Ac_y7_K16K12.0Ac_763}, \text{MS2_G4R17_2Ac_y7_K16K12.1Ac_760}, \text{MS2_G4R17_2Ac_y7_K16K12.2Ac_757})} \right) \Big)
\end{aligned}$$

$$\begin{aligned}
\text{H4.K12acK16ac} = & \frac{\text{MS1_G4R17_2Ac}}{\text{sum}(\text{MS1_G4R17_noPTM}, \text{MS1_G4R17_1Ac}, \text{MS1_G4R17_2Ac}, \text{MS1_G4R17_3Ac}, \text{MS1_G4R17_4Ac})} * \\
& \frac{\text{MS2_G4R17_2Ac_y7_K16K12.2Ac_757}}{\text{sum}(\text{MS2_G4R17_2Ac_y7_K16K12.0Ac_763}, \text{MS2_G4R17_2Ac_y7_K16K12.1Ac_760}, \text{MS2_G4R17_2Ac_y7_K16K12.2Ac_757})}
\end{aligned}$$

$$H4.K5acK8acK12ac = \frac{MS1_G4R17_3Ac}{sum(MS1_G4R17_noPTM, MS1_G4R17_1Ac, MS1_G4R17_2Ac, MS1_G4R17_3Ac, MS1_G4R17_4Ac)} * \left(1 - \frac{MS2_G4R17_3Ac_y5_K16ac_530}{sum(MS2_G4R17_3Ac_y5_K16ac_530, MS2_G4R17_3Ac_y5_K16NoAc_533)} \right)$$

$$H4.K5acK8acK16ac = \frac{MS1_G4R17_3Ac}{sum(MS1_G4R17_noPTM, MS1_G4R17_1Ac, MS1_G4R17_2Ac, MS1_G4R17_3Ac, MS1_G4R17_4Ac)} * \\ \left(\frac{MS2_G4R17_3Ac_y5_K16ac_530}{sum(MS2_G4R17_3Ac_y5_K16ac_530, MS2_G4R17_3Ac_y5_K16NoAc_533)} - \frac{MS2_G4R17_3Ac_y7_K16K12.2Ac_757}{sum(MS2_G4R17_3Ac_y7_K16K12.1Ac_757, MS2_G4R17_3Ac_y7_K16K12.2Ac_760)} \right)$$

$$H4.K5acK12acK16ac = \frac{MS1_G4R17_3Ac}{sum(MS1_G4R17_noPTM, MS1_G4R17_1Ac, MS1_G4R17_2Ac, MS1_G4R17_3Ac, MS1_G4R17_4Ac)} * \\ \left(\frac{MS2_G4R17_3Ac_y7_K16K12.2Ac_757}{sum(MS2_G4R17_3Ac_y7_K16K12.1Ac_757, MS2_G4R17_3Ac_y7_K16K12.2Ac_760)} - \frac{MS2_G4R17_3Ac_y12_K16K12K8.3Ac_1211}{sum(MS2_G4R17_3Ac_y12_K16K12K8.3Ac_1211, MS2_G4R17_3Ac_y12_K16K12K8.2Ac_1214)} \right)$$

$$H4.K8acK12acK16ac = \frac{MS1_G4R17_3Ac}{sum(MS1_G4R17_noPTM, MS1_G4R17_1Ac, MS1_G4R17_2Ac, MS1_G4R17_3Ac, MS1_G4R17_4Ac)} * \frac{MS2_G4R17_3Ac_y12_K16K12K8.3Ac_1211}{sum(MS2_G4R17_3Ac_y12_K16K12K8.3Ac_1211, MS2_G4R17_3Ac_y12_K16K12K8.2Ac_1214)}$$

$$H4.K5acK8acK12acK16ac = \frac{MS1_G4R17_4Ac}{sum(MS1_G4R17_noPTM, MS1_G4R17_1Ac, MS1_G4R17_2Ac, MS1_G4R17_3Ac, MS1_G4R17_4Ac)}$$

Motifs at histone H3

H3.K9R19_noPTM =

$$\frac{\text{MS1_K9R17_noPTM}}{\text{sum}(\text{MS1_K9R17_noPTM}, \text{MS1_K9R17_1Ac}, \text{MS1_K9R17_2Ac}, \text{MS1_K9R17_K9me3K14noAc}, \text{MS1_K9R17_K9me3K14ac}, \text{MS1_K9R17_K9me2K14noAc}, \text{MS1_K9R17_K9me2K14ac}, \text{MS1_K9R17_K9me1K14noAc}, \text{MS1_K9R17_K9me1K14ac})}$$

H3.K9ac =

$$\frac{\text{MS1_K9R17_1Ac}}{\text{sum}(\text{MS1_K9R17_noPTM}, \text{MS1_K9R17_1Ac}, \text{MS1_K9R17_2Ac}, \text{MS1_K9R17_K9me3K14noAc}, \text{MS1_K9R17_K9me3K14ac}, \text{MS1_K9R17_K9me2K14noAc}, \text{MS1_K9R17_K9me2K14ac}, \text{MS1_K9R17_K9me1K14noAc}, \text{MS1_K9R17_K9me1K14ac})}$$

$$\ast \frac{\text{MS1K9R17y7_K14noAc_731}}{\text{sum}(\text{MS1K9R17y7_K14noAc_731}, \text{MS1K9R17y7_K14ac_728})}$$

H3.K14ac =

$$\frac{\text{MS1_K9R17_1Ac}}{\text{sum}(\text{MS1_K9R17_noPTM}, \text{MS1_K9R17_1Ac}, \text{MS1_K9R17_2Ac}, \text{MS1_K9R17_K9me3K14noAc}, \text{MS1_K9R17_K9me3K14ac}, \text{MS1_K9R17_K9me2K14noAc}, \text{MS1_K9R17_K9me2K14ac}, \text{MS1_K9R17_K9me1K14noAc}, \text{MS1_K9R17_K9me1K14ac})}$$

$$\ast \frac{\text{MS1K9R17y7_K14ac_728}}{\text{sum}(\text{MS1K9R17y7_K14noAc_731}, \text{MS1K9R17y7_K14ac_728})}$$

H3.K9me1K14ac =

$$\frac{\text{MS1_K9R17_K9me1K14ac}}{\text{sum}(\text{MS1_K9R17_noPTM}, \text{MS1_K9R17_1Ac}, \text{MS1_K9R17_2Ac}, \text{MS1_K9R17_K9me3K14noAc}, \text{MS1_K9R17_K9me3K14ac}, \text{MS1_K9R17_K9me2K14noAc}, \text{MS1_K9R17_K9me2K14ac}, \text{MS1_K9R17_K9me1K14noAc}, \text{MS1_K9R17_K9me1K14ac})}$$

H3.K9me2K14ac =

$$\frac{\text{MS1_K9R17_K9me2K14ac}}{\text{sum}(\text{MS1_K9R17_noPTM}, \text{MS1_K9R17_1Ac}, \text{MS1_K9R17_2Ac}, \text{MS1_K9R17_K9me3K14noAc}, \text{MS1_K9R17_K9me3K14ac}, \text{MS1_K9R17_K9me2K14noAc}, \text{MS1_K9R17_K9me2K14ac}, \text{MS1_K9R17_K9me1K14noAc}, \text{MS1_K9R17_K9me1K14ac})}$$

H3.K9me3K14ac =

$$\frac{\text{MS1_K9R17_K9me3K14ac}}{\text{sum}(\text{MS1_K9R17_noPTM}, \text{MS1_K9R17_1Ac}, \text{MS1_K9R17_2Ac}, \text{MS1_K9R17_K9me3K14noAc}, \text{MS1_K9R17_K9me3K14ac}, \text{MS1_K9R17_K9me2K14noAc}, \text{MS1_K9R17_K9me2K14ac}, \text{MS1_K9R17_K9me1K14noAc}, \text{MS1_K9R17_K9me1K14ac})}$$

H3.K9me1 =

$$\frac{MS1_K9R17_K9me1K14NoAc}{sum(MS1_K9R17_noPTM, MS1_K9R17_1Ac, MS1_K9R17_2Ac, MS1_K9R17_K9me3K14noAc, MS1_K9R17_K9me3K14ac, MS1_K9R17_K9me2K14noAc, MS1_K9R17_K9me2K14ac, MS1_K9R17_K9me1K14noAc, MS1_K9R17_K9me1K14ac)}$$

H3.K9me2 =

$$\frac{MS1_K9R17_K9me2K14NoAc}{sum(MS1_K9R17_noPTM, MS1_K9R17_1Ac, MS1_K9R17_2Ac, MS1_K9R17_K9me3K14noAc, MS1_K9R17_K9me3K14ac, MS1_K9R17_K9me2K14noAc, MS1_K9R17_K9me2K14ac, MS1_K9R17_K9me1K14noAc, MS1_K9R17_K9me1K14ac)}$$

H3.K9me3 =

$$\frac{MS1_K9R17_K9me3K14NoAc}{sum(MS1_K9R17_noPTM, MS1_K9R17_1Ac, MS1_K9R17_2Ac, MS1_K9R17_K9me3K14noAc, MS1_K9R17_K9me3K14ac, MS1_K9R17_K9me2K14noAc, MS1_K9R17_K9me2K14ac, MS1_K9R17_K9me1K14noAc, MS1_K9R17_K9me1K14ac)}$$

$$H3.K18R26_noPTM = \frac{MS1_K18R26_noPTM}{sum(MS1_K18R26_noPTM, MS1_K18R26_1Ac, MS1_K18R26_2Ac)}$$

$$H3.K18ac = \frac{MS1_K18R26_noPTM}{sum(MS1_K18R26_noPTM, MS1_K18R26_1Ac, MS1_K18R26_2Ac)} * \frac{MS1_K18R26_b2_K18ac_299}{sum(MS1_K18R26_b2_K18ac_299, MS1_K18R26_b2_K18NoAc_302)}$$

$$H3.K23ac = \frac{MS1_K18R26_noPTM}{sum(MS1_K18R26_noPTM, MS1_K18R26_1Ac, MS1_K18R26_2Ac)} * \frac{MS1_K18R26_b2_K18NoAc_302}{sum(MS1_K18R26_b2_K18ac_299, MS1_K18R26_b2_K18NoAc_302)}$$

$$H3.K18acK23ac = \frac{MS1_K18R26_2Ac}{sum(MS1_K18R26_noPTM, MS1_K18R26_1Ac, MS1_K18R26_2Ac)}$$

$$H3.K27R40_noPTM = \frac{MS1_K27R40_noPTM}{sum(MS1_K27R40_noPTM, MS1_K27R40_1Ac, MS1_K27R40_K27me1, MS1_K27R40_K27me2, MS1_K27R40_K27me3, MS1_K27R40_K36me2, MS1_K27R40_K36me3, MS1_K27R40_K27me2K36me1)}$$

$$\frac{MS1_K27R40_noPTM}{MS1_K27R40_K27me2K36ac, MS1_K27R40_K36me2K27ac, MS1_K27R40_K36me3K27ac)}$$

$$\begin{aligned}
 \text{H3.K27ac} = & \frac{\text{MS1_K27R40_1Ac}}{\text{sum}(\text{MS1_K27R40_noPTM}, \text{MS1_K27R40_1Ac}, \text{MS1_K27R40_K27me1}, \text{MS1_K27R40_K27me2}, \text{MS1_K27R40_K27me3}, \text{MS1_K27R40_K36me2}, \text{MS1_K27R40_K36me3}, \text{MS1_K27R40_K27me2K36me1})} \\
 & \frac{\text{MS1_K27R40_noPTM}}{\text{MS1_K27R40_K27me2K36ac}, \text{MS1_K27R40_K36me2K27ac}, \text{MS1_K27R40_K36me3K27ac}} * \\
 & (1 - \text{average}((\frac{\text{MS1_K27R40_y5_K27noAc_752}}{\text{MS1_K27R40_y5_K27noAc_752}, \text{MS1_K27R40_y5_K27ac_755}}), (\frac{\text{MS1_K27R40_b3_K27noAc_332}}{\text{MS1_K27R40_b3_K27noAc_332}, \text{MS1_K27R40_b3_K27ac_329}})))
 \end{aligned}$$

$$\begin{aligned}
 \text{H3.K36ac} = & \frac{\text{MS1_K27R40_1Ac}}{\text{sum}(\text{MS1_K27R40_noPTM}, \text{MS1_K27R40_1Ac}, \text{MS1_K27R40_K27me1}, \text{MS1_K27R40_K27me2}, \text{MS1_K27R40_K27me3}, \text{MS1_K27R40_K36me2}, \text{MS1_K27R40_K36me3}, \text{MS1_K27R40_K27me2K36me1})} \\
 & \frac{\text{MS1_K27R40_noPTM}}{\text{MS1_K27R40_K27me2K36ac}, \text{MS1_K27R40_K36me2K27ac}, \text{MS1_K27R40_K36me3K27ac}} * \\
 & \text{average}((\frac{\text{MS1_K27R40_y5_K27noAc_752}}{\text{MS1_K27R40_y5_K27noAc_752}, \text{MS1_K27R40_y5_K27ac_755}}), (\frac{\text{MS1_K27R40_b3_K27noAc_332}}{\text{MS1_K27R40_b3_K27noAc_332}, \text{MS1_K27R40_b3_K27ac_329}}))) \\
 & * (\frac{\text{MS1_K27R40_y4_K36ac_582}}{\text{MS1_K27R40_y4_K36ac_579}, \text{MS1_K27R40_y4_K36NoAc_582}})
 \end{aligned}$$

$$\begin{aligned}
 \text{H3.K37ac} = & \frac{\text{MS1_K27R40_1Ac}}{\text{sum}(\text{MS1_K27R40_noPTM}, \text{MS1_K27R40_1Ac}, \text{MS1_K27R40_K27me1}, \text{MS1_K27R40_K27me2}, \text{MS1_K27R40_K27me3}, \text{MS1_K27R40_K36me2}, \text{MS1_K27R40_K36me3}, \text{MS1_K27R40_K27me2K36me1})} \\
 & \frac{\text{MS1_K27R40_noPTM}}{\text{MS1_K27R40_K27me2K36ac}, \text{MS1_K27R40_K36me2K27ac}, \text{MS1_K27R40_K36me3K27ac}} * \\
 & \text{average}((\frac{\text{MS1_K27R40_y5_K27noAc_752}}{\text{MS1_K27R40_y5_K27noAc_752}, \text{MS1_K27R40_y5_K27ac_755}}), (\frac{\text{MS1_K27R40_b3_K27noAc_332}}{\text{MS1_K27R40_b3_K27noAc_332}, \text{MS1_K27R40_b3_K27ac_329}}))) \\
 & * (1 - \frac{\text{MS1_K27R40_y4_K36ac_582}}{\text{MS1_K27R40_y4_K36ac_579}, \text{MS1_K27R40_y4_K36NoAc_582}})
 \end{aligned}$$

$$H3.K27R40_K27me1 = \frac{MS1_K27R40_K27me1}{sum(MS1_K27R40_noPTM, MS1_K27R40_1Ac, MS1_K27R40_K27me1, MS1_K27R40_K27me2, MS1_K27R40_K27me3, MS1_K27R40_K36me2, MS1_K27R40_K36me3, MS1_K27R40_K27me2K36me1, MS1_K27R40_K27me2K36ac, MS1_K27R40_K36me2K27ac, MS1_K27R40_K36me3K27ac)}$$

$$H3.K27R40_K27me2 = \frac{MS1_K27R40_K27me2}{sum(MS1_K27R40_noPTM, MS1_K27R40_1Ac, MS1_K27R40_K27me1, MS1_K27R40_K27me2, MS1_K27R40_K27me3, MS1_K27R40_K36me2, MS1_K27R40_K36me3, MS1_K27R40_K27me2K36me1, MS1_K27R40_K27me2K36ac, MS1_K27R40_K36me2K27ac, MS1_K27R40_K36me3K27ac)}$$

$$H3.K27R40_K27me3 = \frac{MS1_K27R40_K27me3}{sum(MS1_K27R40_noPTM, MS1_K27R40_1Ac, MS1_K27R40_K27me1, MS1_K27R40_K27me2, MS1_K27R40_K27me3, MS1_K27R40_K36me2, MS1_K27R40_K36me3, MS1_K27R40_K27me2K36me1, MS1_K27R40_K27me2K36ac, MS1_K27R40_K36me2K27ac, MS1_K27R40_K36me3K27ac)}$$

$$H3.K27R40_K36me2 = \frac{MS1_K27R40_K36me2}{sum(MS1_K27R40_noPTM, MS1_K27R40_1Ac, MS1_K27R40_K27me1, MS1_K27R40_K27me2, MS1_K27R40_K27me3, MS1_K27R40_K36me2, MS1_K27R40_K36me3, MS1_K27R40_K27me2K36me1, MS1_K27R40_K27me2K36ac, MS1_K27R40_K36me2K27ac, MS1_K27R40_K36me3K27ac)}$$

$$H3.K27R40_K36me3 = \frac{MS1_K27R40_K36me3}{sum(MS1_K27R40_noPTM, MS1_K27R40_1Ac, MS1_K27R40_K27me1, MS1_K27R40_K27me2, MS1_K27R40_K27me3, MS1_K27R40_K36me2, MS1_K27R40_K36me3, MS1_K27R40_K27me2K36me1, MS1_K27R40_K27me2K36ac, MS1_K27R40_K36me2K27ac, MS1_K27R40_K36me3K27ac)}$$

$$H3.K27R40_K27me2K36me1 = \frac{MS1_K27R40_K27me2K36me1}{sum(MS1_K27R40_noPTM, MS1_K27R40_1Ac, MS1_K27R40_K27me1, MS1_K27R40_K27me2, MS1_K27R40_K27me3, MS1_K27R40_K36me2, MS1_K27R40_K36me3, MS1_K27R40_K27me2K36me1, MS1_K27R40_K27me2K36ac, MS1_K27R40_K36me2K27ac, MS1_K27R40_K36me3K27ac)}$$

H3.K27R40_K27me2K36ac =

$$\frac{\text{MS1_K27R40_K27me2K36ac}}{\text{sum}(\text{MS1_K27R40_noPTM}, \text{MS1_K27R40_1Ac}, \text{MS1_K27R40_K27me1}, \text{MS1_K27R40_K27me2}, \text{MS1_K27R40_K27me3}, \text{MS1_K27R40_K36me2}, \text{MS1_K27R40_K36me3}, \text{MS1_K27R40_K27me2K36me1})}$$

$$\frac{\text{MS1_K27R40_noPTM}}{\text{MS1_K27R40_K27me2K36ac}, \text{MS1_K27R40_K36me2K27ac}, \text{MS1_K27R40_K36me3K27ac}})$$

H3.K27R40_K36me2K27ac =

$$\frac{\text{MS1_K27R40_K36me2K27ac}}{\text{sum}(\text{MS1_K27R40_noPTM}, \text{MS1_K27R40_1Ac}, \text{MS1_K27R40_K27me1}, \text{MS1_K27R40_K27me2}, \text{MS1_K27R40_K27me3}, \text{MS1_K27R40_K36me2}, \text{MS1_K27R40_K36me3}, \text{MS1_K27R40_K27me2K36me1})}$$

$$\frac{\text{MS1_K27R40_noPTM}}{\text{MS1_K27R40_K27me2K36ac}, \text{MS1_K27R40_K36me2K27ac}, \text{MS1_K27R40_K36me3K27ac}})$$

H3.K27R40_K36me3K27ac =

$$\frac{\text{MS1_K27R40_K36me3K27ac}}{\text{sum}(\text{MS1_K27R40_noPTM}, \text{MS1_K27R40_1Ac}, \text{MS1_K27R40_K27me1}, \text{MS1_K27R40_K27me2}, \text{MS1_K27R40_K27me3}, \text{MS1_K27R40_K36me2}, \text{MS1_K27R40_K36me3}, \text{MS1_K27R40_K27me2K36me1})}$$

$$\frac{\text{MS1_K27R40_noPTM}}{\text{MS1_K27R40_K27me2K36ac}, \text{MS1_K27R40_K36me2K27ac}, \text{MS1_K27R40_K36me3K27ac}})$$

$$\text{H3.E73R83_noPTM} = \frac{\text{MS1_E73R83_noPTM}}{\text{sum}(\text{MS1_E73R83_noPTM}, \text{MS1_E73R83_K79me1}, \text{MS1_E73R83_K79me2})}$$

$$\text{H3.E73R83.K79me1} = \frac{\text{MS1_E73R83_K79me1}}{\text{sum}(\text{MS1_E73R83_noPTM}, \text{MS1_E73R83_K79me1}, \text{MS1_E73R83_K79me2})}$$

$$\text{H3.E73R83.K79me2} = \frac{\text{MS1_E73R83_K79me2}}{\text{sum}(\text{MS1_E73R83_noPTM}, \text{MS1_E73R83_K79me1}, \text{MS1_E73R83_K79me2})}$$

Supplementary Table 3: LC-MS response correction factor (LCF)

H3 motif	LCF (applied)	LCF (measured)	LCF (prior to Q tag normali-sation)	comment
H3.K9K14_noPTM	1.5	ND	ND	estimation of applied LCF based on averaged K9ac/K14ac
H3.K9me1	1.0	1.00 ± 0.12	1.00 ± 0.27	
H3.K9me2	7.0	6.84 ± 0.76	11.44 ± 4.00	
H3.K9me3	34.0	34.17 ± 4.28	50.49 ± 5.82	
H3.K9ac	1.5	1.46 ± 0.48	3.92 ± 0.71	
H3.K14ac	1.5	1.46 ± 0.17	1.62 ± 0.35	
H3.K9acK14ac	1.5	1.58 ± 0.17	2.26 ± 0.31	
H3.K18K23d3_noPTM	1.0	ND	ND	estimation of applied LCF based on averaged K18ac/K23ac
H3.K18ac	1.0	1.23 ± 0.22	0.54 ± 0.05	
H3.K23ac	1.0	1.00 ± 0.09	1.00 ± 0.48	
H3.K18acK23ac	1.5	1.43 ± 0.24	0.51 ± 0.01	
H3.K27d3_noPTM	1	ND	ND	
H3.K27me1	11	11.13 ± 2.49	15.84 ± 20.615	low quality Q tag signal
H3.K27me2	3	3.33 ± 0.61	1.79 ± 1.65	
H3.K27me3	4	4.18 ± 0.12	2.64 ± 2.36	
H3.K27ac	1	ND	ND	
H3.K36me1	1	1 ± 0.25	1.00 ± 0.20	
H3.K36me2	1	1.06 ± 0.73	5.91 ± 0.69	
H3.K36me3	1	0.89 ± 0.02	1.51 ± 0.37	low quality Q tag signal
H3.K27me2K36me1	2	2.34 ± 1.03	7.82 ± 1.68	
H3.K79d3_noPTM	1	ND	ND	
H3.K79me1	2	2.01 ± 0.03	1.49 ± 0.12	
H3.K79me2	1	1.00 ± 0.03	1.00 ± 0.09	
H3.K79me3	2	2.02 ± 0.00	1.67 ± 0.10	

Note: Shown are mean values from replicate experiments. Where applicable, SEM is shown.

Table S4: Histone PTM inventory for KC and S2 cells

motif	% abun- dance KC prior to LCF	% abun- dance KC	Estimated molecules / diploid female cell	% abun- dance S2	Estimated molecules / diploid male cell	% CV (tech.)	% CV (bio.)
H4.K5ac	3.44	3.44 ± 0.08	2.1E+05	2.45 ± 0.13	1.3E+05	1.92	6.66
H4.K8ac	3.09	3.09 ± 0.20	1.8E+05	3.15 ± 0.37	1.6E+05	5.50	8.59
H4.K12ac	7.27	7.27 ± 0.35	4.3E+05	6.24 ± 0.04	3.2E+05	2.23	6.70
H4.K16ac	3.29	3.29 ± 0.12	2.0E+05	8.44 ± 0.06	4.4E+05	0.70	4.52
H4.K5acK8ac	0.59	0.59 ± 0.05	3.5E+04	0.58 ± 0.05	3.0E+04	3.41	3.41
H4.K5acK12ac	4.50	4.50 ± 0.30	2.7E+05	2.89 ± 0.23	1.5E+05	3.51	10.88
H4.K8acK12ac	0.80	0.80 ± 0.09	4.8E+04	1.14 ± 0.45	5.9E+04	22.23	35.73
H4.K8acK16ac	0.41	0.41 ± 0.07	2.5E+04	0.58 ± 0.35	3.0E+04	25.09	24.66
H4.K12acK16ac	0.79	0.79 ± 0.04	4.7E+04	1.48 ± 0.12	7.7E+04	5.16	5.66
H4.K5acK8ac K12ac	0.62	0.62 ± 0.06	3.7E+04	0.55 ± 0.04	2.8E+04	6.71	9.90
H4.K5acK8ac K16ac	0.13	0.13 ± 0.01	7.6E+03	0.19 ± 0.01	9.6E+03	19.24	17.15
H4.K5acK12ac K16ac	0.25	0.25 ± 0.03	1.5E+03	0.45 ± 0.04	2.3E+04	10.40	30.02
H4.K8acK12ac K16ac	0.32	0.32 ± 0.04	1.9E+04	0.49 ± 0.04	2.5E+04	1.72	15.78
H4.K5acK8ac K12acK16ac	0.30	0.30 ± 0.06	1.8E+04	0.57 ± 0.09	2.9E+04	21.58	9.66
H3.K9ac	0.96	0.41 ± 0.04	2.5E+04	0.21 ± 0.01	1.1E+04	11.43	11.23
H3.K14ac	25.63	10.98 ± 0.98	6.6E+05	6.55 ± 0.24	3.4E+05	14.30	1.79
H3.K9acK14ac	1.10	0.47 ± 0.06	2.8E+04	0.27 ± 0.01	1.4E+04	16.86	12.94

H3.K9me1K14ac	5.52	1.60 ± 0.21	9.5E+04	1.49 ± 0.08	7.7E+04	12.98	7.99
H3.K9me2K14ac	1.93	3.72 ± 0.33	2.2E+05	5.02 ± 0.14	2.6E+05	19.82	7.69
H3.K9me3K14ac	0.48	4.44 ± 0.31	2.6E+05	5.48 ± 0.84	2.8E+05	25.98	12.15
H3.K18ac	2.41	2.34 ± 0.09	1.4E+05	2.71 ± 0.24	1.4E+05	2.18	4.00
H3.K23ac	48.56	47.19 ± 0.83	2.8E+06	44.99 ± 0.74	2.3E+06	1.25	3.48
H3.K18acK23ac	5.76	8.40 ± 0.28	5.0E+05	8.62 ± 0.41	4.5E+05	3.96	4.42
H3.K27ac	0.69	0.20 ± 0.03	1.2E+04	0.25 ± 0.04	1.3E+04	3.02	13.12
H3.K36ac	0.20	0.058 ± 0.01	3.5E+03	0.09 ± 0.002	4.8E+03	7.42	10.98
H3.K37ac	0.011	0.003 ± 0.0004	1.82E+02	0.005 ± 0.001	2.5E+02	17.53	18.80
H3.K27ac K36me2	0.09	0.026 ± 0.004	1.5E+03	0.03 ± 0.01	1.4E+03	21.97	13.99
H3.K27ac K36me3	0.26	0.075 ± 0.003	4.5E+03	0.06 ± 0.001	3.1E+03	2.44	26.60
H3.K27me2 K36ac	0.79	0.22 ± 0.05	1.3E+04	0.55 ± 0.04	2.8E+04	11.61	16.46
H3.K9me1	10.42	3.00 ± 0.35	1.8E+05	3.56 ± 0.27	1.8E+05	10.85	10.22
H3.K9me2	10.31	19.98 ± 1.63	1.2E+06	20.01 ± 0.31	1.0E+06	33.79	8.06
H3.K9me3	4.18	38.60 ± 2.16	2.3E+06	44.9 ± 1.75	2.3E+06	35.81	8.81

H3.K27me1	7.14	22.53 ± 1.82	1.3E+06	21.97 ± 0.67	1.1E+06	2.89	8.69
H3.K27me2	41.69	36.25 ± 1.63	2.2E+06	32.47 ± 0.17	1.7E+06	0.88	5.32
H3.K27me3	29.29	33.88 ± 0.93	2.0E+06	36.88 ± 0.71	1.9E+06	0.34	6.16
H3.K36me2	1.18	0.34 ± 0.03	2.0E+04	0.61 ± 0.01	3.2E+04	4.85	11.68
H3.K36me3	3.81	1.10 ± 0.08	6.6E+04	1.55 ± 0.01	8.0E+04	4.67	9.66
H3.K27me2 K36me1	3.59	2.05 ± 0.33	1.2E+05	2.65 ± 0.12	1.4E+04	20.54	11.76
H3.K79me1	20.12	33.46 ± 0.86	2.0E+06	35.67 ± 1.29	1.8E+06	3.92	7.50
H3.K79me2	8.73	7.26 ± 0.49	4.3E+05	6.66 ± 0.06	3.4E+05	2.62	14.55
H4.G4R17 noPTM	74.04	74.04 ± 1.20	4.4E+06	70.17 ± 0.95	3.6E+06	0.29	0.80
H3.K9R17 noPTM	39.48	16.80 ± 1.19	1.0E+06	12.51 ± 0.71	6.5E+05	12.24	8.86
H3.K14R26 noPTM	43.27	42.06 ± 1.10	2.5E+06	43.68 ± 0.84	2.3E+06	0.66	3.99
H3.K27R40 noPTM	11.20	3.25 ± 0.34	1.9E+05	2.89 ± 0.09	1.5E+05	0.99	11.33
H3.E73R83 noPTM	71.15	59.28 ± 1.14	3.5E+06	57.66 ± 1.23	3.0E+06	3.71	4.81

Notes

- % CV for technical replicates (column 6) was derived from 3 whole-workflow replicates using 2 million KC cells
- % CV for biological replicates (column 7) was derived from 5 biological replicates (using 'C' batch control RNAi dataset)
- KC dataset: average of 6 biological 'replicate batches' (batch A-E and technical whole-workflow replicate); S2 dataset: average of 3 biological replicates. '±' indicates SEM. For estimation of cellular copy number: see Suppl. Note 1.3.

Supplementary Table 6: List of *Drosophila melanogaster* KATs and KDACs

Putative KATs analysed in this study

Name used in this study	Nomenclature by Allis et al.	Synonyms (<i>D. mel</i> , human)
HAT1	KAT1	
GCN5	KAT2	hPCAF (KAT2A), hGCN5 (KAT2B)
CBP	KAT3	dNEJIRE (CG15319), hCBP (KAT3A), hP300 (KAT3B)
TAF1	KAT4	TAF250
TIP60	KAT5	PLIP (splice variant)
KAT6	KAT6	dENOK, KAT6A (MYST3/MOZ), KAT6B (MYST4/MORF)
HBO1	KAT7	dCHAMEAU, MYST2
MOF	KAT8	MYST1
ELP3	KAT9	
ATAC2	KAT14*	hCSR2BP
NAT6		CG8481, hFUS2
NAT9		CG11539
NAT10		CG1994, hALP, hKRE33
NAA10		CG11989 (ARD1, VNC)
NAA20		CG14222
NAA30		CG11412
NAA40		CG7593
NAA50		SAN (CG12352), NAT5, MAK3
NAA60		HAT4
CG5783		
CG12560		
ECO		ESCO1/2
MGEA5		OGA (CG5871), MEA5, NCOAT

Putative KDACs analysed in this study

Name used in this study	Nomenclature by Allis et al	Synonyms (<i>D. mel</i> , human)
RPD3		CG7471, hHDAC1/2
HDAC3		CG2128
HDAC4		CG1770, hHDAC4/5/7/9
HDAC6		dHDAC2 (CG6170), hHDAC10
SIR2		CG5216, hSIRT1
SIRT2		CG5085
SIRT4		CG3187
SIRT6		CG6284

Putative KATs not analysed in this study

Name	Comments
CLOCK (CG7391)	restricted expression pattern (preferentially late developmental stages), not expressed in KC cells
CG31493	restricted expression pattern (male accessory gland), not expressed in KC
CG31730	restricted expression pattern (testis-specific), not expressed in KC
CG31851	restricted expression pattern (testis-specific), not expressed in KC
CG32319	restricted expression pattern (testis-specific), not expressed in KC
CG1894	restricted expression pattern, not expressed in KC cells
CG17681	restricted expression pattern, not expressed in KC cells
CG9486	restricted expression pattern, not expressed in KC cells
CG15766	restricted expression pattern, not expressed in KC cells
CG18607	restricted expression pattern, not expressed in KC cells
CG34010	restricted expression pattern, not expressed in KC cells
CG15628	restricted expression pattern, not expressed in KC cells
CG14615	restricted expression pattern, not expressed in KC cells
CG31248	restricted expression pattern, not expressed in KC cells
CG10659	restricted expression pattern, not expressed in KC cells
CG10476	restricted expression pattern, not expressed in KC cells
CG18606	restricted expression pattern, not expressed in KC cells
CG13759	restricted expression pattern, not expressed in KC cells
DAT1 (CG3318)	Dopamine N acetyltransferase
SAT1 (CG4210)	spermidine/spermine N1-acetyltransferase 1
GNPNAT1 (CG1969)	glucosamine-phosphate N-acetyltransferase 1

Putative KDACs not analysed in this study

Name	Comments
HDAC11 (dHDAC-X, CG31119)	restricted expression pattern, not expressed in KC cells
SIRT7 (CG11305)	restricted expression pattern, not expressed in KC cells

Notes

- Column 2 indicates new nomenclature to histone lysine acetyltransferases according to: Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhata, R., et al. (2007). New nomenclature for chromatin-modifying enzymes. Cell 131, 633-636.
- The name KAT14 was introduced after the initial publication by Allis et al. 2008 (Suganuma, T., Gutierrez, J.L., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M., and Workman, J.L. (2008). ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. Nature structural & molecular biology 15, 364-372.)

Supplementary Table 7: List of primers used in this study**Primer used for RNAi**

primer name	primer sequence	comment
GST_for	TTAATACGACTCACTATAGGGAGAATGTCCCC TATACTAG GTTA	Feller et al. 2012
GST_rev	TTAATACGACTCACTATAGGGAGAACGCAT CCAGGCACATTG	
GFP_for	TTAATACGACTCACTATAGGGTGCTCAGGTA GTGGTTGTCG	Regnard et al. 2011
GFP_rev	TTAATACGACTCACTATAGGGCCTGAAGTTCA TCTGCACCA	
EGFP_for	TTAATACGACTCACTATAGGGAGAACGTAAACGGC CACAAGTTCAGC	
EGFP_rev	TTAATACGACTCACTATAGGGAGATGCTGGTAGTG GTCGGCGAG	
HAT1_1_for	TAATACGACTCACTATAGGGACGAAAAGTACA AGAACAACGA	DRSC12298
HAT1_1_rev	TAATACGACTCACTATAGGGATCCATTTTCTT CAAATCCTTTAAC	
HAT1_2_for	TAATACGACTCACTATAGGGACATCTACTTGG GCGTCGATT	DRSC29235
HAT1_2_rev	TAATACGACTCACTATAGGGATCCACGAACC ACAAGATGAA	
HAT1_3_for	TAATACGACTCACTATAGGGGTATTCACCGTC TTGCGTTG	Next-RNAi
HAT1_3_rev	TAATACGACTCACTATAGGGATGAAAGGTCA GGGCATCAT	
GCN5_1_for	TAATACGACTCACTATAGGGAGCGAGATTCT GAAGGAGCTG	DRSC26932
GCN5_1_rev	TAATACGACTCACTATAGGGACACTGCATCAT TTGTCCAC	
GCN5_2_for	TAATACGACTCACTATAGGGAATGAAATCGGT GTCCGAGTC	DRSC35936
GCN5_2_rev	TAATACGACTCACTATAGGGAATGTAGCCCG CATAAACTGG	
CBP_1_for	TAATACGACTCACTATAGGGACGGTGGAAAA GTTGTCTGCT	DRSC25087
CBP_1_rev	TAATACGACTCACTATAGGGAAATTCCTTTGC TAATTGCCG	
CBP_2_for	TAATACGACTCACTATAGGGAGTGGCAGTTA GCCAACAGGT	DRSC26774
CBP_2_rev	TAATACGACTCACTATAGGGAATTGGTCGATT CATTACGGG	
TAF1_1_for	TAATACGACTCACTATAGGGAGAACCAGAAG CTTCAGCCAG	DRSC36664
TAF1_1_rev	TAATACGACTCACTATAGGGATGCTGAGCGT TACTTTGGTG	
TAF1_2_for	TAATACGACTCACTATAGGGAGTCTGGGACG GTAACGACAT	DRSC36665
TAF1_2_rev	TAATACGACTCACTATAGGGACGTTGGGATC CAAAGTGAGT	
TIP60_1_for	TTAATACGACTCACTATAGGGACCTACAACCTC CGCAGACA	DRSC18661
TIP60_1_rev	TTAATACGACTCACTATAGGGAGAAGCTCTGTC ATTACGTAGAAG	

TIP60_2_for	TTAATACGACTCACTATAGGGATTGGCAACTG AAAAGAAGGG	DRSC31332
TIP60_2_rev	TTAATACGACTCACTATAGGGAAAATGTTGAT GACATGGGCA	
TIP60_3_for	TAATACGACTCACTATAGGGCTTTATCAGCCA AAACCCGA	DRSC37670
TIP60_3_rev	TAATACGACTCACTATAGGGAAGCTACAACG AAGCCAGGA	
KAT6_1_for	TTTAATACGACTCACTATAGGGACGTCTAATG AGGCAAAGAAAC	DRSC04096
KAT6_1_rev	TTTAATACGACTCACTATAGGGACCGTTTTTG CCACTTTAACC	
KAT6_2_for	TTAATACGACTCACTATAGGGAGTATCGACTT GTCGCCCAAT	DRSC32736
KAT6_1_rev	TTAATACGACTCACTATAGGGAGGTGACGAG CATGTAGAGCA	
HBO1_1_for	TAATACGACTCACTATAGGGACATGCGAAGG ACAAAACAG	DRSC03341
HBO1_1_rev	TAATACGACTCACTATAGGGATTGTTATTCGC CGGCTTAA	
HBO1_2_for	TAATACGACTCACTATAGGGAGCTACAAAAT GGCCTTGGA	DRSC24941
HBO1_2_rev	TAATACGACTCACTATAGGGAGTTTTCTCCTG GCTGCTGAG	
MOF_1_for	TTAATACGACTCACTATAGGGAGAATGTCTGA AGCGGAGCTGGAAC	Straub et al. 2005
MOF_1_rev	TTAATACGACTCACTATAGGGAGATTTCTGCT TCTGCGGCTGC	
MOF_2_for	TAATACGACTCACTATAGGGCTATGACTACAG CGACCGCA	
MOF_2_rev	TAATACGACTCACTATAGGGCAGCGACTTCTT TTCCTTGG	
ELP3_1_for	TAATACGACTCACTATAGGGAACAGCAGCGC CAATGTG	DRSC00448
ELP3_1_rev	TAATACGACTCACTATAGGGATGGTACAAGG GCCAAAATCT	
ELP3_2_for	TAATACGACTCACTATAGGGACGGCCTAAAG ATCTATCCCA	Next-RNAi
ELP3_2_rev	TAATACGACTCACTATAGGGCGCACCTTGTT GTGGATT	
ATAC_1_for	TTAATACGACTCACTATAGGGAGTTTGAAAG CGATGATTTTAC	DRSC02053
ATAC_1_rev	TTAATACGACTCACTATAGGGACTAGGAATG GCCATGTCATC	
ATAC_2_for	TTAATACGACTCACTATAGGGACGCAAGAGG CAAACGAAAGCA	Next-RNAi
ATAC_2_rev	TTAATACGACTCACTATAGGGAGCCGCACAT AGGCAGGTATGGA	
NAT6_1_for	TAATACGACTCACTATAGGGACCTAATCAACG CAGAGTGG	DRSC16429
NAT6_1_rev	TAATACGACTCACTATAGGGAGCGTTCGTAG AATCCATCC	
NAT6_2_for	TAATACGACTCACTATAGGGATATCTGTCCAC CATCGACCA	DRSC28038
NAT6_2_rev	TAATACGACTCACTATAGGGAGAGGCAATCA CACATGCACT	
NAT9_1_for	TAATACGACTCACTATAGGGAGCCCAAGTAC CACGAGTG	DRSC14325
NAT9_1_rev	TAATACGACTCACTATAGGGACAGCCAGTTG ATCCAATCTG	

NAT10_1_for	TAATACGACTCACTATAGGGGTATTTGGTATT CATGGCCAGT	DRSC18198
NAT10_1_rev	TAATACGACTCACTATAGGGCCTTCTTGCCCC TGCCC	
NAT10_2_for	TAATACGACTCACTATAGGGTGCTGAAGAAC AAGTCCGTG	DRSC35809
NAT10_2_rev	TAATACGACTCACTATAGGGAAAATTTGCCCA ACAACTGG	
NAA10_1_for	TAATACGACTCACTATAGGGACCATGCAGCA CTGCAATC	DRSC09867
NAA10_1_rev	TAATACGACTCACTATAGGGACCATGTCCGC CGGATCT	
NAA10_2_for	TAATACGACTCACTATAGGGGCTGCCCTGAA CCTCTACAC	Next-RNAi
NAA10_2_rev	TAATACGACTCACTATAGGGAATGGCCTTCAA ACCCTTTA	
NAA20_1_for	TAATACGACTCACTATAGGGAACCACGTTGC GACCCTT	DRSC19579
NAA20_1_rev	TAATACGACTCACTATAGGGAGAAATGTCCTC CAGGAAACTC	
NAA20_2_for	TAATACGACTCACTATAGGGGAGAAAGGCCT TATCCAGGG	DRSC38508
NAA20_2_rev	TAATACGACTCACTATAGGGGTGGTTATTTAC GCATCGGC	
NAA30_1_for	TAATACGACTCACTATAGGGACCAAAAAGTGTC GGAGGATGT	DRSC23214
NAA30_1_rev	TAATACGACTCACTATAGGGACCACGTA CTG ATTGTCGTGC	
NAA30_2_for	TAATACGACTCACTATAGGGAAATCGAGATC GGATTGATCG	DRSC30080
NAA30_2_rev	TAATACGACTCACTATAGGGATTTGGATTGTG ATTGGGGTT	
NAA40_1_for	TAATACGACTCACTATAGGGAAAGAATCCCCT CGAATCTCT	DRSC16265
NAA40_1_rev	TAATACGACTCACTATAGGGAGATCTCGTCCT TGACGTAG	
NAA50_1_for	TAATACGACTCACTATAGGGAAGCAGCATCG AACTGGG	DRSC07695
NAA50_1_rev	TAATACGACTCACTATAGGGATGTCGAAGTTT CCGTCCTT	
NAA50_2_for	TAATACGACTCACTATAGGGAAACGATAACTT GCGACCTGC	DRSC31317
NAA50_2_rev	TAATACGACTCACTATAGGGACTGCAGTCAC AATCCTCCAA	
NAA60_1_for	TAATACGACTCACTATAGGGAAGCTGTGCCA AGAATGGTT	DRSC10301
NAA60_1_rev	TAATACGACTCACTATAGGGACTCTTCTCGTA GAAAAATATGG	
NAA60_2_for	TAATACGACTCACTATAGGGAATTCCACCACC TCTGTCTGC	DRSC26416
NAA60_2_rev	TAATACGACTCACTATAGGGAAGCTTGTGGTA GAACCACCG	
NAA60_3_for	TAATACGACTCACTATAGGGGCAAAGAACAA CATCAAAGGC	Next-RNAi
NAA60_3_rev	TAATACGACTCACTATAGGGTAGAAACAGCG AACGGGATT	
CG5783_1_for	TAATACGACTCACTATAGGGAACAAGGCGGC TACTTGG	DRSC02894
CG5783_1_rev	TAATACGACTCACTATAGGGACTCGGGCCAT GTGAACTG	

CG5783_2_for	TAATACGACTCACTATAGGGAGCCTTGTTGG AAAAGCTCTG	DRSC35394
CG5783_2_rev	TAATACGACTCACTATAGGGAAATCCAACACT GACGCACAG	
CG12560_1_for	TAATACGACTCACTATAGGGAGATTGGATCC CTGGGTCTG	DRSC02174
CG12560_1_rev	TAATACGACTCACTATAGGGAGCAGCCCAGT ATGTATTGTC	
ECO_1_for	TAATACGACTCACTATAGGGACAACCTCCCG TCCACTATT	DRSC11006
ECO_1_rev	TAATACGACTCACTATAGGGACGCTGGCCCCG GATGTT	
ECO_2_for	TAATACGACTCACTATAGGGACGGTCCAGAA ATCCCCAAC	VDRC13980
ECO_2_rev	TAATACGACTCACTATAGGGAGGTCGTCTCC TCCGCATCAA	
MGEA5_1_for	TAATACGACTCACTATAGGGACAGCCCCAGG TTTCTCTTC	DRSC15901
MGEA5_1_rev	TAATACGACTCACTATAGGGAGAGAGCCAAG CGTGGGT	
MGEA5_2_for	TAATACGACTCACTATAGGGAGGACGAACAG CAAGTCAAC	derivative of VDRC10644
MGEA5_2_rev	TAATACGACTCACTATAGGGATGGAGGGTGG CGAGGAGGT	
RPD3_1_for	TTAATACGACTCACTATAGGGAGGAGGAGGC GTTCTATACC	DRSC08696
RPD3_1_rev	TTAATACGACTCACTATAGGGAGACTAATGTG CAGCTTAAAATC	
RPD3_2_for	TTAATACGACTCACTATAGGGATCCTCTTCGC GAATTTGTCT	DRSC32186
RPD3_2_rev	TTAATACGACTCACTATAGGGACGCTTTTTGC TGTGAGACTG	
RPD3_3_for	TAATACGACTCACTATAGGGCGCCAAGGAGA ACAACATTT	DRSC37634
RPD3_3_rev	TAATACGACTCACTATAGGGCCACATTGGATC GCTTGTA	
HDAC3_1_for	TTAATACGACTCACTATAGGGACGGGCGGAC TGCATCA	DRSC12352
HDAC3_1_rev	TTAATACGACTCACTATAGGGAACCGCCTCC GACCACC	
HDAC3_2_for	TTAATACGACTCACTATAGGGACGAGGATTTG GTCCACAAC	DRSC32474
HDAC3_2_rev	TTAATACGACTCACTATAGGGATCATTTTCTT GTGCAGTCCG	
HDAC3_3_for	TAATACGACTCACTATAGGGCTTCCACAGCG ACGAGTACA	DRSC32475
HDAC3_3_rev	TAATACGACTCACTATAGGGTGGTTGTGGTTC AGCTTCTG	
HDAC4_1_for	TAATACGACTCACTATAGGGATGAGCACACTA TACGCCAGC	DRSC25237
HDAC4_1_rev	TAATACGACTCACTATAGGGAAAATCGATAAC ACAACCGGC	
HDAC4_2_for	TAATACGACTCACTATAGGGAATATCGCTGCA TTCCGTACC	DRSC37980
HDAC4_2_rev	TAATACGACTCACTATAGGGAGTTGATCGTCT CGGACTCGT	
HDAC6_1_for	TAATACGACTCACTATAGGGAGGATAATCGTT TGCCTGGAA	DRSC29348
HDAC6_1_rev	TAATACGACTCACTATAGGGATCGCTGAGCG TTTTTACCTT	

HDAC6_2_for	TAATACGACTCACTATAGGGAATGACGCTGC TGGGATCTAC	DRSC38396
HDAC6_2_rev	TAATACGACTCACTATAGGGAATGTTACATT GAAGCCCCCT	
SIR2_1_for	TAATACGACTCACTATAGGGAGGATGAGTAC CACACGGTC	DRSC03435
SIR2_1_rev	TAATACGACTCACTATAGGGAAACGCTTTCCA CAGCTAAAA	
SIR2_2_for	TAATACGACTCACTATAGGGAAGGATGAGGA CTCCAGCTCA	DRSC36016
SIR2_2_rev	TAATACGACTCACTATAGGGAGCTCCCGTTA GCACAATGAT	
SIRT2_1_for	TAATACGACTCACTATAGGGATCTGGACTACT TCGAAAAGAA	DRSC15714
SIRT2_1_rev	TAATACGACTCACTATAGGGACCACTGCATC GCGATTG	
SIRT2_2_for	TAATACGACTCACTATAGGGAAGTACGAGCT GCCACATCCT	DRSC26964
SIRT2_2_rev	TAATACGACTCACTATAGGGACAGATCGCAAT CCTGGAAAT	
SIRT4_1_for	TAATACGACTCACTATAGGGAAAGCGTGCG CCTCTAC	DRSC18290
SIRT4_1_rev	TAATACGACTCACTATAGGGATCCAGCGGTA TCTCTACATC	
SIRT4_2_for	TAATACGACTCACTATAGGGAGGCTATGTGG TCAAGTGCCT	DRSC27472
SIRT4_2_rev	TAATACGACTCACTATAGGGACGGAGAAGAC CAGGAGACTG	
SIRT6_1_for	TAATACGACTCACTATAGGGAATTGCACGGC AACATTTACA	DRSC29965
SIRT6_1_rev	TAATACGACTCACTATAGGGATTGCTTGTGG GTATTGTCCA	
NSL1_1_for	TTAATACGACTCACTATAGGGAGCGTCCGAG CTCAACCTTC	Feller et al. 2012
NSL1_1_rev	TTAATACGACTCACTATAGGGACACATGGGT GTGTTCATTAGTC	Feller et al. 2012
NSL1_2_for	TTAATACGACTCACTATAGGGAGATGTCGCAT CAAAGTCAGAGG	Feller et al. 2012
NSL1_2_rev	TTAATACGACTCACTATAGGGAGACTCGAGA AGAGCTCGCTGAT	Feller et al. 2012
NSL1_3_for	TTAATACGACTCACTATAGGG AGATGGCCCCAGCGCTCACA	Raja et al. 2010
NSL1_3_rev	TTAATACGACTCACTATAGGGAGATGAACTTG TGGCCACTGCC	Raja et al. 2010
MSL1_for	TTAATACGACTCACTATAGGGAGAATGGAC AAGCGATTCAAG	
MSL1_rev	TTAATACGACTCACTATAGGGAGACTTCGC TGGTTCTTTTCG	
HeLa_luciferase	CUUACGCUGAGUACUUCGA	gift from Sandra Hake, LMU Munich
HeLa_control1	AAUUCUCCGAACGUGUCAC	gift from Sandra Hake, LMU Munich
hMOF1	GUGAUCCAGUCUCGAGUGA	Taipale et al. 2005
hMOF2	UGCUGUACAGAAGAACUC	Smith et al. 2005
hMOF3	GGAAAGAGAUCUACCGCAA	designed using Dharmagon siDesignCenter

Primer used for RT-qPCR

primer name	primer sequence
gapdh_cDNA_for	GTGACCTACGCAGAAAGCTAG
gapdh_cDNA_rev	GCTATTACGACTGCCGCTTTTTC
HAT1_1_cDNA_for	CTTTCAGAAACTCGGCTTG
HAT1_1_cDNA_rev	TCCTCAGATGGATCCTCCAC
HAT1_2_cDNA_for	TATTCACCGTCTTGCGTTGT
HAT1_2_cDNA_rev	CTTTTTGCAGGCGTTCTAGC
GCN5_cDNA_for	GCCTGCGGATGTAGTAATGC
GCN5_cDNA_rev	CGCGACACATTAACCGGAAA
CBP_1_cDNA_for	GCGGCGGTAAGGATAACATA
CBP_1_cDNA_rev	TGGGTTTTATGTCCACAGCA
TAF1_1_cDNA_for	GAGGTAACGATGGCAAGGAA
TAF1_1_cDNA_rev	TTGCGAACTGCTTGATGAAC
TAF1_2_cDNA_for	GGGCGTTAAAAGGGAGGAT
TAF1_2_cDNA_rev	GTCCTCGGATAACTCGGTGA
TIP60_1_cDNA_for	TGTCTGATCTGGGTTTGCTG
TIP60_1_cDNA_rev	CGGGTTTTGGCTGATAAAGA
TIP60_2_cDNA_for	TTTCTCACCTATCCGCAAG
TIP60_2_cDNA_rev	GCACTTCCTGCTCTTCCTGT
KAT6_cDNA_for	TTGCGGAGACAGTATTGCAG
KAT6_cDNA_rev	TCCGGCGAAAGCTTTAAGTA
HBO1_1_cDNA_for	AGATATACCGCAAGGGCAAA
HBO1_1_cDNA_rev	CAGGCACAGATGCTGACAGT
HBO1_2_cDNA_for	AGCCGAGCTCATCTCCATAA
HBO1_2_cDNA_rev	AGCCAAGAATTTCGTCATCGT
MOF_1_cDNA_for	AGAAGGAGGACGGTCAGGAT
MOF_1_cDNA_rev	TCCACCACCTCTTTCTGCTT
MOF_2_cDNA_for	CTGGGTAGGCTGAGCTATCG
MOF_2_cDNA_rev	TTTCGCTCAGCTCCTTGATT
ELP3_1_cDNA_for	AGTGCATTGGCATCACCATA
ELP3_1_cDNA_rev	CTTCGTAAACCGATTGCACA
ELP3_2_cDNA_for	TTCGGTATGTTGCTGATGGA
ELP3_2_cDNA_rev	TAGTTTCTGGTGCCCACTCC
ATAC2_1_cDNA_for	ATGACATGGCCATTCTAGC
ATAC2_1_cDNA_rev	CTGATGCTCCAGTGCGAATA
ATAC2_2_cDNA_for	AAAGGATACCTTGGCTGCTG
ATAC2_2_cDNA_rev	AGAATCCATGGTGTCCCAAG
NAT6_cDNA_for	TTTGTGCTCTCACACGGAAC
NAT6_cDNA_rev	CGACGAAGTGAACAGAACGA
NAT9_cDNA_for	GGAGATCTTCCACGAGGTCA
NAT9_cDNA_rev	CACTGCATACGCAGATCCAC
NAT10_1_cDNA_for	ACGCCCTGTTTTCGTATCAC
NAT10_1_cDNA_rev	TGCAAATCGTTTGGTGTGTT
NAT10_2_cDNA_for	CAGTCCTCGGACGAACCTAC
NAT10_2_cDNA_rev	AACCGTTTCCACAGTTCGAG
NAA10_1_cDNA_for	TCTGGAAGGCGTTTCTATCG

NAA10_1_cDNA_rev	GATGATCCCAAGCAAGTGGT
NAA10_2_cDNA_for	TTAGAATGTCCCTGCAAACG
NAA10_2_cDNA_rev	CAAAATAATTGCTGCGACGA
NAA20_1_cDNA_for	GATCTGCCCGGTGATAACAT
NAA20_1_cDNA_rev	GGGAAATAATGCACGCAAAT
NAA20_2_cDNA_for	CTATTGCCGGCGAGTATAGC
NAA20_2_cDNA_rev	GTCGCAACGTGGTCATCTTA
NAA30_cDNA_for	AAGCCCATAACAGCCACAAG
NAA30_cDNA_rev	CACAGCTACATCCTCCGACA
NAA40_1_cDNA_for	AAGAACTGGGCTCGCTACTT
NAA40_1_cDNA_rev	GATCCATGTCTGAAGCGGAAC
NAA40_2_cDNA_for	GGCTACGTCAAGGACGAGAT
NAA40_2_cDNA_rev	TCAGCTTAGCATCGATTTGC
NAA50_1_cDNA_for	CGGCGCCTGTATATCATGAC
NAA50_1_cDNA_rev	CGGCGAAGTTCATAATGTGC
NAA50_2_cDNA_for	CGCGAACTCAAATTCACGTA
NAA50_2_cDNA_rev	GCCGCCTCTACATGTCTGTT
NAA60_1_cDNA_for	AATCGTTGCCGAAATCAAAC
NAA60_1_cDNA_rev	CTTGTCTGCATGGGAAAAC
NAA60_2_cDNA_for	CAAAAACACGCACCAAACCTG
NAA60_2_cDNA_rev	TTAAGCGCGAAATCTTGTCC
CG5783_1_cDNA_for	TTGGCAGTGTGGTGACTAGG
CG5783_1_cDNA_rev	CATCTCCGATGACGGCTTAT
CG5783_2_cDNA_for	TACAAGCAAAACTGGCCAAA
CG5783_2_cDNA_rev	ATTCCTCATATGGGGCTGTT
CG12560_1_cDNA_for	GAGCGAAAATTTGGATCGAG
CG12560_1_cDNA_rev	CGCGTAGAAAACCCAAAGAA
CG12560_2_cDNA_for	ATACGGATCCCCAAGGCTAT
CG12560_2_cDNA_rev	AGTCGCCATTCAACGATAGG
ECO_1_cDNA_for	GTGGAGAAGCGGTCACAAAG
ECO_1_cDNA_rev	AATAGTGGACGGGAGGTTGG
ECO_2_cDNA_for	GAGATCGCTAGGGAGTGCAT
ECO_2_cDNA_rev	GAAGTTATCCAACCCGGTGA
MGEA5_1_cDNA_for	GTACACGGAAGTGCCTGAGA
MGEA5_1_cDNA_rev	AAGGGAGGTGACGAGCTGTA
MGEA5_2_cDNA_for	CCTGTTCTATCTGCCCTTCG
MGEA5_2_cDNA_rev	ACATTAGCGTTGCCCTTCAG
RPD3_1_cDNA_for	GTGCCCATTATCAGCAAGGT
RPD3_1_cDNA_rev	TGACGGTGAGATTGAAGCAG
RPD3_2_cDNA_for	GCGAAGATTGTCCCGTCTTC
RPD3_2_cDNA_rev	ATTCAGTTTTACGGCCGCAG
HDAC3_1_cDNA_for	CACGGGAGACATGTACGAGA
HDAC3_1_cDNA_rev	TGGTCATCAATGCCCTCCTT
HDAC3_2_cDNA_for	GATGGCCTCTTCGACTTCTG
HDAC3_2_cDNA_rev	TATGTCGCTGTGGTTGTGGT
HDAC4_1_cDNA_for	CGGACATTGTGCTGGTATCC
HDAC4_1_cDNA_rev	GTCATGAACCCAAAGCAGGC

HDAC4_2_cDNA_for	AATAGAAGGCGATGCAGACG
HDAC4_2_cDNA_rev	TCCCACTCTACCAGGTCCAC
HDAC6_1_cDNA_for	ATTTCGGGATTTTCGGTCTGGA
HDAC6_1_cDNA_rev	TGGA CT CGAAAATGTGCTGC
HDAC6_2_cDNA_for	CGCCACTAAGGCACCTAATC
HDAC6_2_cDNA_rev	CTCAGTGGGCCAGGAAGTAG
SIR2_1_cDNA_for	CTGTGCTGCGATGAGTCAGT
SIR2_1_cDNA_rev	GTGGCGGTGGTGAAGTAGAT
SIR2_2_cDNA_for	ATACACTCGCACCCACATCA
SIR2_2_cDNA_rev	ATTTCAAACACGCACACCAA
SIRT2_1_cDNA_for	GGATTTTCAGATCCCCAGGTT
SIRT2_1_cDNA_rev	GATCGAATATGGCCGTAGGA
SIRT2_2_cDNA_for	TGTCAACTTCGGGACATGAA
SIRT2_2_cDNA_rev	GCAGCTGAAACTGCAACAAA
SIRT4_1_cDNA_for	TGAAGCCGGAAATTGTCTTC
SIRT4_1_cDNA_rev	GGCCATCGCTATTGTACACC
SIRT4_2_cDNA_for	CGTGCGCTCACCTCTAGTAA
SIRT4_2_cDNA_rev	CCCACACGCATTGTTTTGT
SIRT6_1_cDNA_for	ATTGCACGGCAACATTTACA
SIRT6_1_cDNA_rev	ACGTTGCAGGGATTTTTGAC
SIRT6_2_cDNA_for	TGTGGATGTGGTTTTGTCCA
SIRT6_2_cDNA_rev	GGCTTGGA CTGCTTTGTAGG

SUPPLEMENTAL REFERENCES

- Becker, P.B., and Wu, C. (1992). Cell-free system for assembly of transcriptionally repressed chromatin from *Drosophila* embryos. *Molecular and cellular biology* 12, 2241-2249.
- Das, C., Lucia, M.S., Hansen, K.C., and Tyler, J.K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* 459, 113-117.
- Di Cerbo, V., Mohn, F., Ryan, D.P., Montellier, E., Kacem, S., Tropberger, P., Kallis, E., Holzner, M., Hoerner, L., Feldmann, A., *et al.* (2014). Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. *eLife* 3, e01632.
- Drogaris, P., Villeneuve, V., Pomies, C., Lee, E.H., Bourdeau, V., Bonneil, E., Ferbeyre, G., Verreault, A., and Thibault, P. (2012). Histone deacetylase inhibitors globally enhance h3/h4 tail acetylation without affecting h3 lysine 56 acetylation. *Scientific reports* 2, 220.
- Farley, A.R., and Link, A.J. (2009). Identification and quantification of protein posttranslational modifications. *Methods in enzymology* 463, 725-763.
- Feller, C., Prestel, M., Hartmann, H., Straub, T., Soding, J., and Becker, P.B. (2012). The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset. *Nucleic acids research* 40, 1509-1522.
- Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., *et al.* (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature genetics* 37, 391-400.
- Gallien, S., Duriez, E., Crone, C., Kellmann, M., Moehring, T., and Domon, B. (2012a). Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Molecular & cellular proteomics : MCP* 11, 1709-1723.
- Gallien, S., Peterman, S., Kiyonami, R., Souady, J., Duriez, E., Schoen, A., and Domon, B. (2012b). Highly multiplexed targeted proteomics using precise control of peptide retention time. *Proteomics* 12, 1122-1133.
- Goodlett, D.R., Keller, A., Watts, J.D., Newitt, R., Yi, E.C., Purvine, S., Eng, J.K., von Haller, P., Aebersold, R., and Kolker, E. (2001). Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation. *Rapid communications in mass spectrometry : RCM* 15, 1214-1221.
- Gu, W., Szauter, P., and Lucchesi, J.C. (1998). Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of *Drosophila melanogaster*. *Developmental genetics* 22, 56-64.
- Jaffé, J.D., Wang, Y., Chan, H.M., Zhang, J., Huether, R., Kryukov, G.V., Bhang, H.E., Taylor, J.E., Hu, M., Englund, N.P., *et al.* (2013). Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nature genetics* 45, 1386-1391.
- Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC genomics* 13, 424.

- Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., *et al.* (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471, 480-485.
- Kueh, A.J., Dixon, M.P., Voss, A.K., and Thomas, T. (2011). HBO1 is required for H3K14 acetylation and normal transcriptional activity during embryonic development. *Molecular and cellular biology* 31, 845-860.
- Lalonde, M.E., Avvakumov, N., Glass, K.C., Joncas, F.H., Saksouk, N., Holliday, M., Paquet, E., Yan, K., Tong, Q., Klein, B.J., *et al.* (2013). Exchange of associated factors directs a switch in HBO1 acetyltransferase histone tail specificity. *Genes & development* 27, 2009-2024.
- Leroy, G., Chepelev, I., Dimaggio, P.A., Blanco, M.A., Zee, B.M., Zhao, K., and Garcia, B.A. (2012). Proteogenomic characterization and mapping of nucleosomes decoded by Brd and HP1 proteins. *Genome biology* 13, R68.
- Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Molecular cell* 24, 309-316.
- Marx, H., Lemeer, S., Schliep, J.E., Matheron, L., Mohammed, S., Cox, J., Mann, M., Heck, A.J., and Kuster, B. (2013). A large synthetic peptide and phosphopeptide reference library for mass spectrometry-based proteomics. *Nature biotechnology* 31, 557-564.
- Makowski, A.M., Dutnall, R.N., and Annunziato, A.T. (2001). Effects of acetylation of histone H4 at lysines 8 and 16 on activity of the Hat1 histone acetyltransferase. *The Journal of biological chemistry* 276, 43499-43502.
- Miotto, B., and Struhl, K. (2010). HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by Geminin. *Molecular cell* 37, 57-66.
- Peterson, A.C., Russell, J.D., Bailey, D.J., Westphall, M.S., and Coon, J.J. (2012). Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Molecular & cellular proteomics : MCP* 11, 1475-1488.
- Phanstiel, D., Brumbaugh, J., Berggren, W.T., Conard, K., Feng, X., Levenstein, M.E., McAlister, G.C., Thomson, J.A., and Coon, J.J. (2008). Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 105, 4093-4098.
- Prestel, M., Feller, C., Straub, T., Mitlohner, H., and Becker, P.B. (2010). The activation potential of MOF is constrained for dosage compensation. *Molecular cell* 38, 815-826.
- Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Analytical chemistry* 75, 663-670.
- Schmidt, A., Gehlenborg, N., Bodenmiller, B., Mueller, L.N., Campbell, D., Mueller, M., Aebersold, R., and Domon, B. (2008). An integrated, directed mass spectrometric approach for in-depth characterization of complex peptide mixtures. *Molecular & cellular proteomics : MCP* 7, 2138-2150.

- Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callen, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I.A., *et al.* (2008). A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes & development* 22, 2048-2061.
- Smith, E.R., Cayrou, C., Huang, R., Lane, W.S., Cote, J., and Lucchesi, J.C. (2005). A human protein complex homologous to the *Drosophila* MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Molecular and cellular biology* 25, 9175-9188.
- Taipale, M., Rea, S., Richter, K., Vilar, A., Lichter, P., Imhof, A., and Akhtar, A. (2005). hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. *Molecular and cellular biology* 25, 6798-6810.
- Tang, H., Fang, H., Yin, E., Brasier, A.R., Sowers, L.C., and Zhang, K. (2014). Multiplexed parallel reaction monitoring targeting histone modifications on the QExactive mass spectrometer. *Analytical chemistry* 86, 5526-5534.
- Thorne, A.W., Kmiecik, D., Mitchelson, K., Sautiere, P., and Crane-Robinson, C. (1990). Patterns of histone acetylation. *European journal of biochemistry / FEBS* 193, 701-713.
- Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E.T., Buhler, M., *et al.* (2013). Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell* 152, 859-872.
- Verreault, A., Kaufman, P.D., Kobayashi, R., and Stillman, B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell* 87, 95-104.
- Yanez-Cuna, J.O., Arnold, C.D., Stampfel, G., Boryn, L.M., Gerlach, D., Rath, M., and Stark, A. (2014). Dissection of thousands of cell type-specific enhancers identifies dinucleotide repeat motifs as general enhancer features. *Genome research* 24, 1147-1156.
- Yang, X., Yu, W., Shi, L., Sun, L., Liang, J., Yi, X., Li, Q., Zhang, Y., Yang, F., Han, X., *et al.* (2011). HAT4, a Golgi apparatus-anchored B-type histone acetyltransferase, acetylates free histone H4 and facilitates chromatin assembly. *Molecular cell* 44, 39-50.
- Ye, J., Ai, X., Eugeni, E.E., Zhang, L., Carpenter, L.R., Jelinek, M.A., Freitas, M.A., and Parthun, M.R. (2005). Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly. *Molecular cell* 18, 123-130.
- Young, N.L., DiMaggio, P.A., Plazas-Mayorca, M.D., Baliban, R.C., Floudas, C.A., and Garcia, B.A. (2009). High throughput characterization of combinatorial histone codes. *Molecular & cellular proteomics : MCP* 8, 2266-2284.
- Zhang, K., Sridhar, V.V., Zhu, J., Kapoor, A., and Zhu, J.K. (2007). Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *PloS one* 2, e1210.
- Zheng, Y., Thomas, P.M., and Kelleher, N.L. (2013). Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites. *Nature communications* 4, 2203.

3.4.4 Discussion on substrate specificity of lysine acetyltransferases using the examples of HAT1 and HBO1

For many lysine acetyltransferases, there is no consensus with regard to their histone target sites. As described before (section 2.3.1), most KATs function in the context of multi-subunit protein complexes, where accessory subunits are often required for a robust KAT activity and specificity towards distinct sites. This strong dependency on co-factors complicates the ability to measure the substrate specificity *in vitro*, which necessitates well-defined reconstituted systems matching closely the *in vivo* observed stoichiometries and protein modifications. The observed specificities vary strongly when using different substrates, such as histones, octamers or nucleosomes and when performing the reactions with different KAT sub-complexes (Grant et al., 1999; Morales et al., 2004; Lee and Workman, 2007).

Previous studies characterizing the *in vivo* specificity to a given KAT often appear controversial. This may be attributed to different cellular contexts or technical limitations. Most of our current knowledge of KAT – substrate relationships is based on characterizations using antibodies. Antibodies raised against specific histone acetylation sites may show substantial lot-to-lot variations and have often been associated with substantial cross-reactivity (Egelhofer et al., 2011; Fuchs et al., 2011; Nishikori et al., 2012; Rothbart et al., 2012). Moreover, in the most cases, only a few putative histone target sites were tested. This hampers the overall comparability between individual studies and makes it difficult to arrive at a comprehensive conclusion. Using the examples of HAT1 and HBO1, I discuss below the reported substrate specificities for these enzymes and how this relates to the data presented in section 3.4.2.

HAT1 not only di-acetylates lysines 5 and 12 on H4 to promote nuclear import of histones but it also contributes to H3 lysine 27 acetylation, which may provide links to its nuclear function on DNA repair, histone turnover and transcription.

Histone acetyltransferase 1 (HAT1) is considered to be the main cytosolic histone lysine acetyltransferase (Verreault et al., 1996) and the only known KAT that generates a combinatorial histone PTM. Di-acetylation on lysines 5 and 12 on the newly synthesised histone H4 molecules by HAT1 promotes their association as H3/H4 dimers with chaperones and their subsequent nuclear import and chromatin assembly (Parthun, 2012). Although newly synthesised histones are marked by H4.K5acK12ac in most species (Ruiz-Carrillo et al., 1975; Chicoine et al., 1986; Sobel et al., 1995), it is less well understood how HAT1 produces this combinatorial mark and whether there are additional HAT1 target sites on histones. Moreover, recent studies link HAT1 to nuclear functions including DNA repair (Parthun, 2012) yet whether this involves different histone targets remains elusive.

In this study, we confirm that HAT1 is the major KAT responsible for H4.K5acK12ac. Moreover, we identify putative novel targets on H4 (K5ac, K5acK8acK12ac) and H3 (K27ac, K27acK36me2). Reduction of mono-acetylated K5 on histone H4 (H4.K5ac) is unexpected, because the majority of *in vitro* studies using recombinant HAT1 identified its main activity towards H4.K12 (Kleff et al., 1995; Parthun et al., 1996; Verreault et al., 1996; Dose et al., 2011). Interestingly, our observation on the reduction of H4.K5ac is well in line with early studies, which showed that the cytoplasmic HAT-B activity purified from the cytoplasm or micronuclei of *Tetrahymena* produces di-acetylated H4 by first acetylating K5 and subsequently K12 (Richman et al., 1988). In addition, our observation of reduced tri-acetylated H4.K5acK8acK12ac is consistent with the finding that the HAT-B activity from pie converts non-modified H4 into mono-acetylated, di-acetylated and tri-acetylated H4 isoforms (Mingarro et al., 1993). Moreover, newly synthesised histones complexed within chaperones or immunoprecipitated with HAT1 contain acetylated K8 but not K16 in addition to acetylated K5 and 12 (Verreault et al., 1996; Yang et al., 2013).

Mechanistic insights into the specificity towards the di-acetylation H4.K5acK12ac by HAT1 came from a series of structural and biochemical studies. Modeling the histone peptides into the crystal structure of HAT1 suggested that lysines 8 and 16 must remain unmodified and positively charged to electrostatically stabilise the acidic patch within the active site of the enzyme (Dutnall et al., 1998). This prediction was verified by a series of *in vitro* acetyltransferase assays comparing a set of pre-acetylated histone peptides as substrates: H4 peptides di-acetylated at lysines 8 and 16 abrogated the KAT activity almost as efficient as di-acetylated peptides at lysines K5 and 12. In contrast, peptides mono-acetylated at K5 or K8 but not K12 qualified as substrates almost as efficient as the unmodified peptide (Makowski et al., 2001; Benson et al., 2007). These mechanistical experiments together with the *in vivo* results presented before (Yang et al., 2013) and our data suggest that after the initial acetylation of H4.K5, HAT1 either releases its substrate (unprocessive mode) or further modifies it to its major product, the di-acetylated H4.K5acK12ac. However, a minor fraction of H4.K5acK12ac is likely not released until it is further processed to tri-acetylated H4.K5acK8acK12ac.

Grunstein and colleagues challenged the paradigm that di-acetylated H4.K5acK12ac is required for nuclear import and proper chromatin assembly. Analysing sets of yeast histone mutants, they concluded that only mutants with lysine to glycine substitutions at three (K5, K8, K12) or four (K5, K8, K12, K16) H4 lysine positions impaired yeast viability and chromatin maturation (Ma et al., 1998). These findings indicate that other histone marks may contribute to the initial setup of chromatin – particular if the default system via H4.K5acK12ac is disturbed. In line, we observe that HAT1 RNAi provokes an increase of the H4 di-acetylation motif H4.K8acK16ac and several H3 acetylation motifs (most of them are not statistically significant, including H3.K9ac, H3.K9me2K14ac and K27acK36me3). It is interesting to note that the H4.K8acK16ac peptide, which efficiently inhibited HAT1 activity *in vitro* (Makowski et al., 2001), is upregulated after HAT1 deprivation. Similarly,

yeast genetic experiments with histone mutants suggested that some phenotypes with H4 tail deletion or lysine point mutations can be rescued by manipulating the charge distribution on histone H3 (Shahbazian and Grunstein, 2007). Moreover, acetylated H3.K14ac and H3.K23ac have been observed at cytoplasmic histones and as part of chaperone complexes (Jackson et al., 1976; Sobel et al., 1995; Tyler et al., 1999). It is tempting to speculate that in the absence of the default H4.K5acK12ac nuclear import signal, a compensatory system uses other H4 di-acetylation motifs or H3 acetylation motifs to serve as a backup strategy.

While the cytoplasmic function of HAT1 is well described, recent attention has been devoted to its nuclear function. HAT1 is pre-dominantly found in the nucleus (Verreault et al., 1998; Poveda et al., 2004; Yang et al., 2013), where it may contribute to ensure proper chromatin assembly by regulating the histone turnover in the wake of the DNA repair machinery (Qin and Parthun, 2002; Verzijlbergen et al., 2011; Yang et al., 2013). Our observation of reduced H3.K27ac, which is considered to be a hallmark of active enhancers, may point to a novel function of HAT1 in transcriptional regulation. This is further supported by the observation that cells depleted for HAT1 show reduced levels of histone methyl marks associated with active transcription (H3.K36me2/3) and an increase of the repressive mark H3.K27me3.

dHBO1/Chameau: The finding that dHBO1 is responsible for most mono-acetylated H4.K12ac and the combinatorial motifs H4.K8acK12ac and H3.K9me3K14ac may contribute to resolve the long-standing controversy whether HBO1 acetylates H3 or H4.

Published reports on HBO1 are controversial with respect to its histone acetylation targets, its role on DNA replication and cell growth and its capacity to promote or repress transcription. Our finding that dHBO1 is responsible for most mono-acetylated H4.K12ac and the combinatorial marks H4.K8acK12ac and H3.K9me3K14ac may contribute to resolve some of the long-standing controversies for HBO1's histone targets. In addition, we did not observe a significant change in the cell cycle or cell number, suggesting that dHBO1 is not required for proper maintenance of the cell cycle, DNA replication and cell growth in *Drosophila* KC cells.

HBO1 was initially described as Histone acetyltransferase Binding to ORC with acetyltransferase activity to histones H3 and H4 (Iizuka and Stillman, 1999). However, most reports following its original description characterised HBO1 as either i) the major acetyltransferase responsible for histone H4 acetylation, ii) the major acetyltransferase for lysine 5, 8 and 12 yet not 16 on histone H4, iii) the major histone H3 acetyltransferase or iv) the major KAT for H3.K14ac (Doyon et al., 2006; Foy et al., 2008; Wu and Liu, 2008; Miotto and Struhl, 2010; Kueh et al., 2011; Mishima et al., 2011; Havasi et al., 2013). Moreover, individual studies reported significant reduction of H3.K9ac or H3.K23ac while the beforehand mentioned sites remain largely unaffected (Hung et al., 2009; Saksouk et al., 2009;

Lalonde et al., 2013). In particular the immuno blot experiments that showed a global loss of histone H4 acetylation using pan-acetyl H4 antibodies are somewhat surprising taking into account that more than 70% of histone H4 acetylation in most human and murine cells is on position lysine 16. Notably, most of the mentioned reports stressed that ablation of HBO1 reduces acetylation on either H3 *or* H4. For example, Kueh *et al* used immuno blotting to demonstrate that *hbo1* knockout mouse primary fibroblasts have greatly diminished H3.K14ac levels, which are accompanied by a strong increase in H4.K16ac and a moderate increase in H4.K5ac and H3.K9ac (Kueh et al., 2011). In contrast, Miotto and Struhl also used immuno blotting yet they reported that HeLa cells depleted for HBO1 show massively decreased acetylation on lysines 5,8 and 12 on H4, while acetylation at H4K16 and H3 stay almost constant (Miotto and Struhl, 2010). Although recently Cote and colleagues suggested that HBO1 co-exists in two different complexes in HeLa cells – where the HBO1-JADE complex acetylates H4 while the HBO1-BRPF1 complex acetylates H3.K23ac (and to a lesser extent H3.K14ac, (Lalonde et al., 2013) – others observed different putative target sites and no concomitant reduction of H3 and H4 acetylation in HeLa cells depleted for HBO1 (Iizuka et al., 2009; Miotto and Struhl, 2010).

To my knowledge, there are only two studies which directly addressed the acetylation target sites of the *Drosophila* homologue to HBO1, Chameau: McConnell and colleagues reported that Chameau and CBP are required together to reduce H4.K12ac in *Drosophila* ovaries, whereas the individual KAT mutants did not affect histone acetylation levels (McConnell et al., 2012). In contrast, Miotto *et al.* reported an increase of H4K16ac on a reporter gene in HEK293 cells when *Drosophila* Chameau is targeted to it (Miotto et al., 2006).

We found that depleting dHBO1 (Chameau) in female KC cells or mutating dHBO1 in male flies (unpublished preliminary results) reduced acetylation of lysine 14 on histone H3 only in the context of adjacently tri-methylated lysine 9 (H3.K9me3K14ac). Notably, while H3.K9me3K14ac contributes to only a minority of all combinations containing acetylated lysine 14 (~ 4%), the major acetylation isoform H3.K14ac (~ 11%) is not reduced but even increased in HBO1 depleted cells. Because, to our knowledge, all reports addressing HBO1 histone targets used anti-acetylation antibodies, modulation of complex epitopes such as the mixed acetylation-methylation motif H3.K9me3K14ac may have skipped previous attention. Importantly, we also observed that dHBO1 depleted cells show a reduction of tri-methylated K9 in the context of unmodified K14 (-17% of H3.K9me3K14). These levels are comparable to the reduction of H3.K9me3K14ac by 31%. These results may indicate that the observed reduction of acetylated K14 could be a secondary effect due to the strong global decrease of K9 tri-methylation. This finding highlights the necessity to also examine non-target modifications at adjacent residues to distinguish putative primary from secondary effects.

Our finding that dHBO1 depleted cells show reduced levels of H4.K12ac and H4.K8acK12ac are in agreement with previous studies which reported that HBO1 preferentially acetylates histone H4 over

H3. However, considering that some antibodies raised against histone H4 acetylation sites recognise multi-acetylated peptides more than an order of magnitude better than mono-acetylated peptides (Rothbart et al., 2012), our finding of reduced H4.K8acK12ac levels may indicate that the previously reported strong reduction of acetylated lysines 5 and 8 upon HBO1 RNAi might be interpreted with caution. This holds in particular true because, at least in *Drosophila* KC cells as well as S2 cells (unpublished results), dCBP is responsible for the majority of mono-acetylated K5 and K8. Finally, our observation that dHBO1 provokes an increase of H4.K16ac and H3.K9ac-containing motifs is in agreement with a previous report (Kueh et al., 2011).

4 GENERAL CONCLUSIONS AND OUTLOOK

Post-translational histone modifications serve as a model to study how chemical modifications diversify protein functions. Because histones and their PTM-derived isoforms are abundant cellular biomolecules they are more easily accessible to experimentation than PTMs at most other proteins. Moreover, because histones are intimately engaged in organizing DNA, individual nucleosomal histone PTMs can be tracked spatiotemporally by determining with which region of DNA they associate. This often allows correlating a specific modification to a biological process. For example, through functional genomic approaches we know of specific individual histone modifications associated with DNA elements such as transcriptional enhancers and promoters.

A further layer of complexity is introduced by the combinatorial potential of co-occurring modifications along the same histone molecule, the partner histones within the same nucleosomes or adjacent nucleosomes. There is accumulating biochemical evidence that pre-set modifications influence the binding and activity of chromatin-modifying activities (Fischle, 2008; Suganuma and Workman, 2011). In some cases, a combinatorial histone PTM motif rather than the individual marks may serve as the biological signal. An illustrative example is the di-acetyl motif H4.K5acK12ac, which is associated with nascent cytosolic histones (Sobel et al., 1995). Moreover, dedicated effector proteins may preferentially recognise combinatorial motifs. Examples include certain bromodomains that show increased affinity towards complex peptide motifs harbouring distinct acetylation patterns (Moriniere et al., 2009; Filippakopoulos et al., 2012) and effector proteins that contain multiple linked histone PTM recognizing domains (Ruthenburg et al., 2007; Patel and Wang, 2013; Rothbart et al., 2013). However, because traditional genomic approaches depend on highly specific antibodies, which are, in most cases, unable to assess complex PTM motifs, new strategies are required to locate combinatorial motifs to functional DNA elements and biological processes. Recent advances in native proximity ligation assays (Koos et al., 2014), multivalent, recombinant and intracellular binder reagents (Haque et al., 2011; Hayashi-Takanaka et al., 2011; Hattori et al., 2013), super-resolution microscopy (Schermelleh et al., 2010; Sauer, 2013), and the mass spectrometry-based analysis of selected genomic regions (Wang et al., 2013a; Pourfarzad et al. 2013; Waldrip et al., 2014) promise attractive experimental strategies to fill some of the current technological gaps. Concepts and technologies developed to analyse combinatorial PTMs at histones are likely to also spur the progress towards understanding complex PTM motifs at other proteins.

One of the next challenges is to study how different types of PTMs cross-talk on a systems level and under varying internal and external conditions, including perturbation scenarios. To what extent does a modification depend on a pre-existing mark? What are the compensatory strategies to circumvent compromising effects if individual PTM-modifying pathways are blocked?

The study presented in section 3.4 provides a first insight in the changes of histone acetylation and methylation patterns in response to systematically depleting every known or suspected KAT and KDAC. However, by analysing the changes after 5.5 days of RNAi mediated protein depletion, we only present a static snapshot of these events. The recent developments of specific inhibitors to KATs and KDACs, acute and reversible protein depletion strategies and the increasing availability of well-characterised constitutive and conditional mutants will enable monitoring the choreography of cellular responses to perturbations at different time scales with potentially fascinating insights between different temporal stages. This will not only allow us to better discriminate primary from secondary effects but will also improve our understanding of how chromatin pathways are interconnected. Previous literature already provided a glimpse in the cross-regulation of individual chromatin-modifying enzymes, as exemplified by their mutual protein-protein interactions (see section 2.3). Monitoring protein and activity levels of a panel of histone PTM modifiers when an individual enzyme is ablated will likely uncover more cross-regulatory potential.

Moreover, we mainly catalogued KATs and KDACs in a single cell line, yet the substrate specificities, functions and cellular responses may vary across cell types, developmental stages and species and may dependent on induced or transient signals or specific environmental conditions. Likewise, we only analysed histone substrates but acetyltransferases also target non-histone proteins. Is there cross-talk between histone and non-histone acetylation in cells deprived of acetyltransferases? Finally, we restricted this study to the analysis of single gene perturbations. It will be of interest how the system of acetylation motifs (and other PTM systems) responds to ablating two or a set of KATs. Will we still detect balanced global histone acetylation levels? What will we learn about redundancy among the enzymes or across histone sites/motifs? Will we uncover local and global epistatic interactions?

Knowledge of the abundance of histone PTMs can facilitate the design and interpretation of genomic and biochemical studies

Quantitative mass spectrometry (MS) studies indicated that the default states of histones in the cell are highly chemically decorated rather than unmodified isoforms. Our study refines this view by using a panel of synthetic peptides to correct for differential MS detection bias and thereby improves the accurateness to quantify histone PTMs. For example, this correction procedure allowed is to describe that a substantial fraction of histone H3 molecules is modified by tri-methylation on lysine 9 (39% instead of 4% without correction), which is in line with the high occurrence of repetitive DNA elements within the *Drosophila* genome that are covered by this histone mark.

Other marks such as the polycomb signatures (H3.K27me2/3) have traditionally been associated with a few developmentally regulated genes yet recent ChIP-chip/Seq studies uncovered in addition domains of substantial sizes (Hawkins et al., 2010). Previous mass spectrometry studies already

demonstrated the high abundance of this mark in mouse cells (40% and 10% of H3.K27me2/3, (Peters et al., 2003)), and we confirmed this finding for *Drosophila* cells (36% and 34% in KC cells).

This concordance between genomic and proteomic studies is not observed for H4.K16ac in mammalian cells. Despite its high abundance, as judged by mass spectrometry (>30%), ChIP-chip/Seq studies describe very limited binding patterns for this mark, preferentially localizing to promoter and some putative enhancer regions (Wang et al., 2008b; Sridharan et al., 2013; Taylor et al., 2013). Thus, the low number of ChIP peaks may be interpreted as high local accumulation of this mark at these loci relative to regions in the direct vicinity. Alternatively, because histone acetylation antibodies have been shown to display a greatly enhanced affinity towards poly-acetylated H4 (Rothbart et al., 2012), the H4.K16ac-ChIP signals may preferentially report on these multi-acetylated H4 isoforms.

In contrast to mammalian cells, H4.K16ac is of lower abundance in *Drosophila* male (8%) and female (3%) cells. Here, ChIP-chip/Seq experiments in female cells identified distinct peaks at promoters of active genes and at intergenic regions associated with replication regions (Kind et al., 2008; Schwaiger et al., 2009). Male cells showed in addition to the female H4.K16ac pattern extensive signals along gene bodies, which is the defining histone signature for fly dosage compensation (Gelbart et al., 2009). Comparing the cellular abundances of this mark between female and male cells as well as male cells with perturbations to the MSL-DCC allows to estimate the density of this modification along the dosage compensated gene bodies. Moreover, employing quantitative ChIP-MS strategies further refines this estimation and likely will uncover new chromatin marks associated with dosage compensation (unpublished results). In summary, knowing the accurate cellular abundances is helpful to design and interpret genomic experiments.

Likewise, information on the cellular abundance of histone PTMs may also guide biochemical studies that are aimed to decipher the effect of selected modifications to the structure of chromatin. Here, knowing the most prevalent patterns of co-occurrence may help to select appropriate chromatin templates to study a selected modification in the context of its naturally co-occurring ‘partner’ modification(s). For example, studying the effect of H4.K16ac in the context of the highly abundant H4.K20me2 mark (Pesavento et al., 2008) may facilitate to resolve some of the current controversies on how ISWI-containing nucleosome remodelers are regulated by this acetylation mark (see 2.3.1). Similar, the rather moderate effects of H3.K56ac and H3.K79me on the nucleosome structure may be modulated in the context of the very prevalent H3.K23ac, H3.K9me3 and H3.K27me2 marks. Recent advances in generating nucleosomal arrays in a fast and streamlined manner, simultaneously characterizing many different templates in quantitative biochemical assays and monitoring the effect of PTM motifs on enzyme kinetics in real-time foreshadow exciting future directions to understand mechanistically how histone modifications modulate chromatin enzymes and structures (Dose et al., 2011; Nguyen et al., 2014).

Feed-forward and feedback as compensatory responses to evolutionary-manifested and spontaneous perturbation

Dosage compensation can be conceptualised as a feed-forward principle, where the system anticipates the genetic perturbation – that is the reduced gene dose caused by the ‘missing’ X chromosome – and hence had evolved a compensatory mechanism to deal with this evolutionary-manifested gene dosage deficiency (for a thorough discussion see section 3.2.2 and (Zhang et al., 2010)). One prediction for such mechanism would be to implement a single activating principle that achieves directly two-fold stimulation on top of the basal gene activity. Indeed, a previous study has shown that tethering MSL2 to a reporter gene brings the expected two-fold activation characteristic for dosage compensation (Straub et al., 2005). In this experiment, MSL2 likely recruits the missing members of the MSL-DCC to reconstitute a full and functionally intact dosage compensation complex.

In our study discussed in section 3.1 we further dissected the ‘activating principle’ by tethering MOF instead of MSL2. Comparing the MOF-driven reporter gene in male and female flies allowed us to recognise that the strong transcriptional activation potential of MOF, which is observable in female flies, is constrained in male flies to the two-fold range. The key experiment was depleting MSL2 in these cells. Under this conditions, transcription from the MOF-driven reporter gene was enhanced rather than reduced, arguing that MSL2 itself or an MSL2-associated factor limits the transcription stimulation brought by MOF. More generally, this result suggests that dosage compensation does not use a single principle to directly achieve ‘two-fold up’, but rather constrains a strong general activator (MOF/H4.K16ac) to arrive at the finely-tuned set point of two-fold activation.

Two recent screens discovered a novel factor that may provide a candidate that constrains the activity of the MSL complex (Larschan et al., 2012; Lim and Kelley, 2013). Employing a genome-wide RNAi screen, Kuroda and colleagues identified CG3363 as the strongest repressor of the MSL-DCC activity on a *roX2* reporter gene (Larschan et al., 2012). In a genetic screen for novel regulators of dosage compensation, Kelley and colleague identified the same gene, which they termed over compensating males (*ocm*), that, when mutated, caused elevated dosage compensation (Lim and Kelley, 2013). This feature is strikingly similar to our observation of increased transcription from the MOF-driven reporter gene after ablating MSL2. Not much is known about the *ocm* gene, which encodes a large, 250 kDa protein without any recognizable functional domain except for a short cysteine rich motif. Future studies addressing the functions of *ocm* are likely to also advance our understanding of balancing activating and repressive principles during fly dosage compensation.

A potentially more indirect way to constrain the strong activation brought about by MOF-mediated hyper-acetylation of H4.K16ac is by modulating heterochromatin components. It has long been known that the male X chromosome is in particular sensitive to the levels of certain heterochromatin proteins. Reducing the dose of NURF, HP1 or the HP1-interactor Su(var)3-7 results in a strong decondensation

phenotype specifically at the male X chromosome ('X bloating') (Spierer et al., 2005; Corona et al., 2007; Spierer et al., 2008). HP1 may directly form more repressive chromatin structures and NURF may promote compaction by incorporation of the linker histone H1 (Corona et al., 2007).

The data presented in section 3.4 provides further support towards an involvement of heterochromatin factors for the process of dosage compensation. Comparing histone modifications between male and female wildtype and MOF-deprived cells allowed us to recognise that male cells contain elevated levels of the heterochromatin marks H3.K9me3 and H3.K27me3. Importantly, ablating MOF (and hence dosage compensation) in male cells re-adjusted these marks towards the levels found in female cells. Extending the current study from analysing bulk histones towards comparing these modifications in cells either containing or lacking MOF using ChIP-Seq will allow determining whether these heterochromatin marks are directly enriched at target genes of the MSL-DCC. In addition, these experiments will allow testing whether the re-distribution of heterochromatin components is an integral part of the feed-forward principle of dosage compensation or whether it is the result of a feedback mechanism.

Many biological systems function as highly connected and responsive networks, where spontaneous perturbation to a single component elicits a feedback mechanism to compensate for the loss of a single member. The most surprising finding of the third study was the observation that depletion of almost every acetyltransferase resulted in a gain of acetylation at secondary sites that balanced the loss of acetylation at primary sites. I speculate that this phenomenon functions as a feedback mechanism to keep the global acetylation levels constant. Testing this hypothesis requires to identify suitable 'compensatory pairs' that allow to quantify the functional consequences after depleting a single or both components of this pair. The system of dosage compensation may provide a suitable model to study this potentially 'compensatory' phenomenon, because of the experimental power to monitor quantitative changes in transcription over thousand active genes, which are all under the regime of the same regulatory system. We observed similar global histone acetylation levels between male and female cells. The strongly increased levels of H4.K16ac in male cells are 'balanced' by reduced levels of H4.K5ac, H4.K12ac and H3.K14ac, and conversely, depleting MOF restores the levels to the ones found in female cells. Comparative ChIP-Seq analysis for these modifications should reveal whether these marks are specifically depleted at dosage compensated genes. Further, manipulating the enzymes that modulate these modifications followed by quantifying the consequence on X-chromosomal gene expression should not only allow testing whether they have a role during dosage compensation but also provide insights into the functional significance of induced acetylation at secondary sites.

5 REFERENCES

- Aka, J.A., Kim, G.W., and Yang, X.J. (2011). K-acetylation and its enzymes: overview and new developments. *Handbook of experimental pharmacology* 206, 1-12.
- Akhtar, A., and Becker, P.B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Molecular cell* 5, 367-375.
- Alekseyenko, A.A., Larschan, E., Lai, W.R., Park, P.J., and Kuroda, M.I. (2006). High-resolution ChIP-chip analysis reveals that the *Drosophila* MSL complex selectively identifies active genes on the male X chromosome. *Genes & development* 20, 848-857.
- Allahverdi, A., Yang, R., Korolev, N., Fan, Y., Davey, C.A., Liu, C.F., and Nordenskiöld, L. (2011). The effects of histone H4 tail acetylations on cation-induced chromatin folding and self-association. *Nucleic acids research* 39, 1680-1691.
- Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 51, 786-794.
- Allfrey, V.G., Littau, V.C., and Mirsky, A.E. (1963). On the role of of histones in regulation ribonucleic acid synthesis in the cell nucleus. *Proceedings of the National Academy of Sciences of the United States of America* 49, 414-421.
- Allfrey, V.G., and Mirsky, A.E. (1962). Evidence for the complete DNA-dependence of RNA synthesis in isolated thymus nuclei. *Proceedings of the National Academy of Sciences of the United States of America* 48, 1590-1596.
- Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhatter, R., *et al.* (2007). New nomenclature for chromatin-modifying enzymes. *Cell* 131, 633-636.
- Amrein, H., and Axel, R. (1997). Genes expressed in neurons of adult male *Drosophila*. *Cell* 88, 459-469.
- Andrews, A.J., and Luger, K. (2011). Nucleosome structure(s) and stability: variations on a theme. *Annual review of biophysics* 40, 99-117.
- Badenhorst, P., Voas, M., Rebay, I., and Wu, C. (2002). Biological functions of the ISWI chromatin remodeling complex NURF. *Genes & development* 16, 3186-3198.
- Badenhorst, P., Xiao, H., Cherbas, L., Kwon, S.Y., Voas, M., Rebay, I., Cherbas, P., and Wu, C. (2005). The *Drosophila* nucleosome remodeling factor NURF is required for Ecdysteroid signaling and metamorphosis. *Genes & development* 19, 2540-2545.
- Baneres, J.L., Essalouh, L., Jariel-Encontre, I., Mesnier, D., Garrod, S., and Parello, J. (1994). Evidence indicating proximity in the nucleosome between the histone H4 N termini and the globular domain of histone H1. *Journal of molecular biology* 243, 48-59.

- Bannister, A.J., and Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature* 384, 641-643.
- Bashaw, G.J., and Baker, B.S. (1997). The regulation of the *Drosophila* msl-2 gene reveals a function for Sex-lethal in translational control. *Cell* 89, 789-798.
- Becker, P.B., and Horz, W. (2002). ATP-dependent nucleosome remodeling. *Annual review of biochemistry* 71, 247-273.
- Bedford, D.C., and Brindle, P.K. (2012). Is histone acetylation the most important physiological function for CBP and p300? *Aging* 4, 247-255.
- Bell, O., Conrad, T., Kind, J., Wirbelauer, C., Akhtar, A., and Schubeler, D. (2008). Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in *Drosophila melanogaster*. *Molecular and cellular biology* 28, 3401-3409.
- Belote, J.M., and Lucchesi, J.C. (1980a). Control of X chromosome transcription by the maleless gene in *Drosophila*. *Nature* 285, 573-575.
- Belote, J.M., and Lucchesi, J.C. (1980b). Male-specific lethal mutations of *Drosophila melanogaster*. *Genetics* 96, 165-186.
- Benson, L.J., Phillips, J.A., Gu, Y., Parthun, M.R., Hoffman, C.S., and Annunziato, A.T. (2007). Properties of the type B histone acetyltransferase Hat1: H4 tail interaction, site preference, and involvement in DNA repair. *The Journal of biological chemistry* 282, 836-842.
- Black, J.C., Van Rechem, C., and Whetstone, J.R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Molecular cell* 48, 491-507.
- Bohm, V., Hieb, A.R., Andrews, A.J., Gansen, A., Rocker, A., Toth, K., Luger, K., and Langowski, J. (2011). Nucleosome accessibility governed by the dimer/tetramer interface. *Nucleic acids research* 39, 3093-3102.
- Bonisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic acids research* 40, 10719-10741.
- Bridges, C.B., Sex in relation to chromosomes and genes. *Am. Nat.*, 1922(56): p. 51-63
- Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843-851.
- Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., *et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581-592.
- Champagne, N., Bertos, N.R., Pelletier, N., Wang, A.H., Vezmar, M., Yang, Y., Heng, H.H., and Yang, X.J. (1999). Identification of a human histone acetyltransferase related to monocytic leukemia zinc finger protein. *The Journal of biological chemistry* 274, 28528-28536.

- Chelmicki, T., Dundar, F., Turley, M.J., Khanam, T., Aktas, T., Ramirez, F., Gendrel, A.V., Wright, P.R., Videm, P., Backofen, R., *et al.* (2014). MOF-associated complexes ensure stem cell identity and Xist repression. *eLife* 3, e02024.
- Chicoine, L.G., Schulman, I.G., Richman, R., Cook, R.G., and Allis, C.D. (1986). Nonrandom utilization of acetylation sites in histones isolated from *Tetrahymena*. Evidence for functionally distinct H4 acetylation sites. *The Journal of biological chemistry* 261, 1071-1076.
- Chodaparambil, J.V., Barbera, A.J., Lu, X., Kaye, K.M., Hansen, J.C., and Luger, K. (2007). A charged and contoured surface on the nucleosome regulates chromatin compaction. *Nature structural & molecular biology* 14, 1105-1107.
- Chopra, V.S., Hendrix, D.A., Core, L.J., Tsui, C., Lis, J.T., and Levine, M. (2011). The polycomb group mutant *esc* leads to augmented levels of paused Pol II in the *Drosophila* embryo. *Molecular cell* 42, 837-844.
- Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. *Annual review of biochemistry* 78, 273-304.
- Clapier, C.R., and Cairns, B.R. (2012). Regulation of ISWI involves inhibitory modules antagonized by nucleosomal epitopes. *Nature* 492, 280-284.
- Clapier, C.R., Langst, G., Corona, D.F., Becker, P.B., and Nightingale, K.P. (2001). Critical role for the histone H4 N terminus in nucleosome remodeling by ISWI. *Molecular and cellular biology* 21, 875-883.
- Conrad, T., and Akhtar, A. (2011). Dosage compensation in *Drosophila melanogaster*: epigenetic fine-tuning of chromosome-wide transcription. *Nature reviews Genetics* 13, 123-134.
- Corona, D.F., Clapier, C.R., Becker, P.B., and Tamkun, J.W. (2002). Modulation of ISWI function by site-specific histone acetylation. *EMBO reports* 3, 242-247.
- Corona, D.F., Siriaco, G., Armstrong, J.A., Snarskaya, N., McClymont, S.A., Scott, M.P., and Tamkun, J.W. (2007). ISWI regulates higher-order chromatin structure and histone H1 assembly in vivo. *PLoS biology* 5, e232.
- Cunliffe, V.T. (2008). Eloquent silence: developmental functions of Class I histone deacetylases. *Current opinion in genetics & development* 18, 404-410.
- Dahlsveen, I.K., Gilfillan, G.D., Shelest, V.I., Lamm, R., and Becker, P.B. (2006). Targeting determinants of dosage compensation in *Drosophila*. *PLoS genetics* 2, e5.
- Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase *Drosophila* cells. *PLoS biology* 5, e218.
- Deardorff, M.A., Bando, M., Nakato, R., Watrin, E., Itoh, T., Minamino, M., Saitoh, K., Komata, M., Katou, Y., Clark, D., *et al.* (2012). HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature* 489, 313-317.

- Deng, X., Berletch, J.B., Ma, W., Nguyen, D.K., Hiatt, J.B., Noble, W.S., Shendure, J., and Distèche, C.M. (2013). Mammalian X upregulation is associated with enhanced transcription initiation, RNA half-life, and MOF-mediated H4K16 acetylation. *Developmental cell* 25, 55-68.
- Deng, X., Berletch, J.B., Nguyen, D.K., and Distèche, C.M. (2014). X chromosome regulation: diverse patterns in development, tissues and disease. *Nature reviews Genetics* 15, 367-378.
- Deuring, R., Fanti, L., Armstrong, J.A., Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S.L., Berloco, M., *et al.* (2000). The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. *Mol Cell* 5, 355-365.
- Devaiah, B.N., and Singer, D.S. (2013). Two faces of brd4: mitotic bookmark and transcriptional lynchpin. *Transcription* 4, 13-17.
- Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399, 491-496.
- Dias, J., Van Nguyen, N., Georgiev, P., Gaub, A., Brettschneider, J., Cusack, S., Kadlec, J., and Akhtar, A. (2014). Structural analysis of the KANSL1/WDR5/KANSL2 complex reveals that WDR5 is required for efficient assembly and chromatin targeting of the NSL complex. *Genes & development* 28, 929-942.
- Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R., and Richmond, T.J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* 306, 1571-1573.
- Dose, A., Liokatis, S., Theillet, F.X., Selenko, P., and Schwarzer, D. (2011). NMR profiling of histone deacetylase and acetyl-transferase activities in real time. *ACS chemical biology* 6, 419-424.
- Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* 121, 873-885.
- Doyon, Y., Cayrou, C., Ullah, M., Landry, A.J., Cote, V., Selleck, W., Lane, W.S., Tan, S., Yang, X.J., and Cote, J. (2006). ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Molecular cell* 21, 51-64.
- Drouin, S., Laramee, L., Jacques, P.E., Forest, A., Bergeron, M., and Robert, F. (2010). DSIF and RNA polymerase II CTD phosphorylation coordinate the recruitment of Rpd3S to actively transcribed genes. *PLoS genetics* 6, e1001173.
- Du, J., Zhou, Y., Su, X., Yu, J.J., Khan, S., Jiang, H., Kim, J., Woo, J., Kim, J.H., Choi, B.H., *et al.* (2011). Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science* 334, 806-809.
- Dupont, C., and Gribnau, J. (2013). Different flavors of X-chromosome inactivation in mammals. *Current opinion in cell biology* 25, 314-321.
- Durant, M., and Pugh, B.F. (2006). Genome-wide relationships between TAF1 and histone acetyltransferases in *Saccharomyces cerevisiae*. *Molecular and cellular biology* 26, 2791-2802.

- Durrin, L.K., Mann, R.K., Kayne, P.S., and Grunstein, M. (1991). Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* 65, 1023-1031.
- Dutnall, R.N., Tafrov, S.T., Sternglanz, R., and Ramakrishnan, V. (1998). Structure of the histone acetyltransferase Hat1: a paradigm for the GCN5-related N-acetyltransferase superfamily. *Cell* 94, 427-438.
- Egelhofer, T.A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A.A., Cheung, M.S., Day, D.S., Gadel, S., Gorchakov, A.A., *et al.* (2011). An assessment of histone-modification antibody quality. *Nature structural & molecular biology* 18, 91-93.
- Feldman, J.L., Dittenhafer-Reed, K.E., and Denu, J.M. (2012). Sirtuin catalysis and regulation. *The Journal of biological chemistry* 287, 42419-42427.
- Feller, C., Prestel, M., Hartmann, H., Straub, T., Soding, J., and Becker, P.B. (2012). The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset. *Nucleic acids research* 40, 1509-1522.
- Ferrari, F., Plachetka, A., Alekseyenko, A.A., Jung, Y.L., Oszolak, F., Kharchenko, P.V., Park, P.J., and Kuroda, M.I. (2013). "Jump start and gain" model for dosage compensation in *Drosophila* based on direct sequencing of nascent transcripts. *Cell reports* 5, 629-636.
- Fierz, B., Chatterjee, C., McGinty, R.K., Bar-Dagan, M., Raleigh, D.P., and Muir, T.W. (2011). Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nature chemical biology* 7, 113-119.
- Filippakopoulos, P., and Knapp, S. (2012). The bromodomain interaction module. *FEBS letters* 586, 2692-2704.
- Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Muller, S., Pawson, T., *et al.* (2012). Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* 149, 214-231.
- Finch, J.T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. *Proceedings of the National Academy of Sciences of the United States of America* 73, 1897-1901.
- Fischle, W. (2008). Talk is cheap--cross-talk in establishment, maintenance, and readout of chromatin modifications. *Genes & development* 22, 3375-3382.
- Fischle, W., Dequiedt, F., Fillion, M., Hendzel, M.J., Voelter, W., and Verdin, E. (2001). Human HDAC7 histone deacetylase activity is associated with HDAC3 in vivo. *The Journal of biological chemistry* 276, 35826-35835.
- Fischle, W., Dequiedt, F., Hendzel, M.J., Guenther, M.G., Lazar, M.A., Voelter, W., and Verdin, E. (2002). Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. *Molecular cell* 9, 45-57.
- Flemming, W. (1882). *Zellsubstanz, Kern und Zellteilung*. Leipzig, F.C.W Vogel.

- Fletcher, T.M., and Hansen, J.C. (1995). Core histone tail domains mediate oligonucleosome folding and nucleosomal DNA organization through distinct molecular mechanisms. *The Journal of biological chemistry* 270, 25359-25362.
- Foy, R.L., Song, I.Y., Chitalia, V.C., Cohen, H.T., Saksouk, N., Cayrou, C., Vaziri, C., Cote, J., and Panchenko, M.V. (2008). Role of Jade-1 in the histone acetyltransferase (HAT) HBO1 complex. *The Journal of biological chemistry* 283, 28817-28826.
- Fuchs, S.M., Krajewski, K., Baker, R.W., Miller, V.L., and Strahl, B.D. (2011). Influence of combinatorial histone modifications on antibody and effector protein recognition. *Current biology : CB* 21, 53-58.
- Fukunaga, A., Tanaka, A., and Oishi, K. (1975). Maleless, a recessive autosomal mutant of *Drosophila melanogaster* that specifically kills male zygotes. *Genetics* 81, 135-141.
- Fullgrabe, J., Lynch-Day, M.A., Heldring, N., Li, W., Struijk, R.B., Ma, Q., Hermanson, O., Rosenfeld, M.G., Klionsky, D.J., and Joseph, B. (2013). The histone H4 lysine 16 acetyltransferase hMOF regulates the outcome of autophagy. *Nature* 500, 468-471.
- Furuyama, T., and Henikoff, S. (2009). Centromeric nucleosomes induce positive DNA supercoils. *Cell* 138, 104-113.
- Fussner, E., Ching, R.W., and Bazett-Jones, D.P. (2011). Living without 30nm chromatin fibers. *Trends in biochemical sciences* 36, 1-6.
- Fyodorov, D.V., Blower, M.D., Karpen, G.H., and Kadonaga, J.T. (2004). Acf1 confers unique activities to ACF/CHRAC and promotes the formation rather than disruption of chromatin in vivo. *Genes & development* 18, 170-183.
- Gelbart, M.E., and Kuroda, M.I. (2009). *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* 136, 1399-1410.
- Gelbart, M.E., Larschan, E., Peng, S., Park, P.J., and Kuroda, M.I. (2009). *Drosophila* MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. *Nature structural & molecular biology* 16, 825-832.
- Gilfillan, G.D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P.B. (2006). Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex. *Genes & development* 20, 858-870.
- Grant, P.A., Eberharter, A., John, S., Cook, R.G., Turner, B.M., and Workman, J.L. (1999). Expanded lysine acetylation specificity of Gcn5 in native complexes. *The Journal of biological chemistry* 274, 5895-5900.
- Gregory, P.D., Schmid, A., Zavari, M., Lui, L., Berger, S.L., and Horz, W. (1998). Absence of Gcn5 HAT activity defines a novel state in the opening of chromatin at the PHO5 promoter in yeast. *Molecular cell* 1, 495-505.

- Grigoryev, S.A., Arya, G., Correll, S., Woodcock, C.L., and Schlick, T. (2009). Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proceedings of the National Academy of Sciences of the United States of America* *106*, 13317-13322.
- Grimaud, C., and Becker, P.B. (2009). The dosage compensation complex shapes the conformation of the X chromosome in *Drosophila*. *Genes & development* *23*, 2490-2495.
- Grimaud, C., and Becker, P.B. (2010). Form and function of dosage-compensated chromosomes--a chicken-and-egg relationship. *BioEssays : news and reviews in molecular, cellular and developmental biology* *32*, 709-717.
- Gu, W., Szauter, P., and Lucchesi, J.C. (1998). Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of *Drosophila melanogaster*. *Developmental genetics* *22*, 56-64.
- Hake, S.B., and Allis, C.D. (2006). Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". *Proceedings of the National Academy of Sciences of the United States of America* *103*, 6428-6435.
- Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. *Cell* *55*, 1137-1145.
- Hansen, J.C., Tse, C., and Wolffe, A.P. (1998). Structure and function of the core histone N-termini: more than meets the eye. *Biochemistry* *37*, 17637-17641.
- Haque, A., Andersen, J.N., Salmeen, A., Barford, D., and Tonks, N.K. (2011). Conformation-sensing antibodies stabilize the oxidized form of PTP1B and inhibit its phosphatase activity. *Cell* *147*, 185-198.
- Hattori, T., Taft, J.M., Swist, K.M., Luo, H., Witt, H., Slattery, M., Koide, A., Ruthenburg, A.J., Krajewski, K., Strahl, B.D., *et al.* (2013). Recombinant antibodies to histone post-translational modifications. *Nature methods* *10*, 992-995.
- Havasi, A., Haegeler, J.A., Gall, J.M., Blackmon, S., Ichimura, T., Bonegio, R.G., and Panchenko, M.V. (2013). Histone acetyl transferase (HAT) HBO1 and JADE1 in epithelial cell regeneration. *The American journal of pathology* *182*, 152-162.
- Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S., *et al.* (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell stem cell* *6*, 479-491.
- Hayashi-Takanaka, Y., Yamagata, K., Wakayama, T., Stasevich, T.J., Kainuma, T., Tsurimoto, T., Tachibana, M., Shinkai, Y., Kurumizaka, H., Nozaki, N., *et al.* (2011). Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. *Nucleic acids research* *39*, 6475-6488.
- Hazzalin, C.A., and Mahadevan, L.C. (2005). Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: analysis at c-fos and c-jun. *PLoS biology* *3*, e393.

- Herrera, J.E., Bergel, M., Yang, X.J., Nakatani, Y., and Bustin, M. (1997). The histone acetyltransferase activity of human GCN5 and PCAF is stabilized by coenzymes. *The Journal of biological chemistry* 272, 27253-27258.
- Herz, H.M., Garruss, A., and Shilatifard, A. (2013). SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends in biochemical sciences* 38, 621-639.
- Herz, H.M., Mohan, M., Garruss, A.S., Liang, K., Takahashi, Y.H., Mickey, K., Voets, O., Verrijzer, C.P., and Shilatifard, A. (2012). Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes & development* 26, 2604-2620.
- Hewish, D.R., and Burgoyne, L.A. (1973). Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochemical and biophysical research communications* 52, 504-510.
- Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J.C. (1997). mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in Drosophila. *The EMBO journal* 16, 2054-2060.
- Hojfeldt, J.W., Agger, K., and Helin, K. (2013). Histone lysine demethylases as targets for anticancer therapy. *Nature reviews Drug discovery* 12, 917-930.
- Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *The Journal of biological chemistry* 268, 305-314.
- Horowitz, R.A., Giannasca, P.J., and Woodcock, C.L. (1990). Ultrastructural preservation of nuclei and chromatin: improvement with low-temperature methods. *Journal of microscopy* 157, 205-224.
- Horowitz, R.A., Koster, A.J., Walz, J., and Woodcock, C.L. (1997). Automated electron microscope tomography of frozen-hydrated chromatin: the irregular three-dimensional zigzag architecture persists in compact, isolated fibers. *Journal of structural biology* 120, 353-362.
- Huang, R.C., and Bonner, J. (1962). Histone, a suppressor of chromosomal RNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 48, 1216-1222.
- Hung, T., Binda, O., Champagne, K.S., Kuo, A.J., Johnson, K., Chang, H.Y., Simon, M.D., Kutateladze, T.G., and Gozani, O. (2009). ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Molecular cell* 33, 248-256.
- Huyen, Y., Zgheib, O., Ditullio, R.A., Jr., Gorgoulis, V.G., Zacharatos, P., Petty, T.J., Sheston, E.A., Mellert, H.S., Stavridi, E.S., and Halazonetis, T.D. (2004). Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* 432, 406-411.
- Iizuka, M., and Stillman, B. (1999). Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. *The Journal of biological chemistry* 274, 23027-23034.

- Iizuka, M., Takahashi, Y., Mizzen, C.A., Cook, R.G., Fujita, M., Allis, C.D., Frierson, H.F., Jr., Fukusato, T., and Smith, M.M. (2009). Histone acetyltransferase Hbo1: catalytic activity, cellular abundance, and links to primary cancers. *Gene* 436, 108-114.
- Imai, S., Armstrong, C.M., Kaeberlein, M., and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* 403, 795-800.
- Jackson, V., Shires, A., Tanphaichitr, N., and Chalkley, R. (1976). Modifications to histones immediately after synthesis. *Journal of molecular biology* 104, 471-483.
- Jeon, Y., Sarma, K., and Lee, J.T. (2012). New and Xisting regulatory mechanisms of X chromosome inactivation. *Current opinion in genetics & development* 22, 62-71.
- Jin, Q., Yu, L.R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.E., Wang, C., Brindle, P.K., Dent, S.Y., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *The EMBO journal* 30, 249-262.
- Jorgensen, S., Schotta, G., and Sorensen, C.S. (2013). Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. *Nucleic acids research* 41, 2797-2806.
- Joshi, P., Greco, T.M., Guise, A.J., Luo, Y., Yu, F., Nesvizhskii, A.I., and Cristea, I.M. (2013). The functional interactome landscape of the human histone deacetylase family. *Molecular systems biology* 9, 672.
- Kan, P.Y., Caterino, T.L., and Hayes, J.J. (2009). The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. *Molecular and cellular biology* 29, 538-546.
- Kasten, M., Szerlong, H., Erdjument-Bromage, H., Tempst, P., Werner, M., and Cairns, B.R. (2004). Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14. *The EMBO journal* 23, 1348-1359.
- Kelley, R.L., Meller, V.H., Gordadze, P.R., Roman, G., Davis, R.L., and Kuroda, M.I. (1999). Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. *Cell* 98, 513-522.
- Kelley, R.L., Solovyeva, I., Lyman, L.M., Richman, R., Solovyev, V., and Kuroda, M.I. (1995). Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila. *Cell* 81, 867-877.
- Kelley, R.L., Wang, J., Bell, L., and Kuroda, M.I. (1997). Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism. *Nature* 387, 195-199.
- Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., *et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 123, 593-605.
- Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., *et al.* (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* 471, 480-485.

- Kim, D., Blus, B.J., Chandra, V., Huang, P., Rastinejad, F., and Khorasanizadeh, S. (2010). Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. *Nature structural & molecular biology* 17, 1027-1029.
- Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P., and Akhtar, A. (2008). Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in *Drosophila*. *Cell* 133, 813-828.
- Kleff, S., Andrulis, E.D., Anderson, C.W., and Sternglanz, R. (1995). Identification of a gene encoding a yeast histone H4 acetyltransferase. *The Journal of biological chemistry* 270, 24674-24677.
- Klinker, H., Mueller-Planitz, F., Yang, R., Forne, I., Liu, C.F., Nordenskiöld, L., and Becker, P.B. (2014). ISWI remodelling of physiological chromatin fibres acetylated at lysine 16 of histone H4. *PloS one* 9, e88411.
- Komma, D.J. (1966). Effect of sex transformation genes on glucose-6-phosphate dehydrogenase activity in *Drosophila melanogaster*. *Genetics* 54, 497-503.
- Koos, B., Andersson, L., Clausson, C.M., Grannas, K., Klaesson, A., Cane, G., and Soderberg, O. (2014). Analysis of protein interactions in situ by proximity ligation assays. *Current topics in microbiology and immunology* 377, 111-126.
- Korber, P. (2012). Active nucleosome positioning beyond intrinsic biophysics is revealed by in vitro reconstitution. *Biochemical Society transactions* 40, 377-382.
- Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868-871.
- Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. *Science* 184, 865-868.
- Korolev, N., Zhao, Y., Allahverdi, A., Eom, K.D., Tam, J.P., and Nordenskiöld, L. (2012). The effect of salt on oligocation-induced chromatin condensation. *Biochemical and biophysical research communications* 418, 205-210.
- Kossel, A. (1884). Ueber einen peptonartigen Bestandtheil des Zellkerns. *Hoppe-Seyler's Zeitschrift für physiologische Chemie*: 511-515.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- Kruithof, M., Chien, F.T., Routh, A., Logie, C., Rhodes, D., and van Noort, J. (2009). Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nature structural & molecular biology* 16, 534-540.
- Kueh, A.J., Dixon, M.P., Voss, A.K., and Thomas, T. (2011). HBO1 is required for H3K14 acetylation and normal transcriptional activity during embryonic development. *Molecular and cellular biology* 31, 845-860.
- Kusch, T., Florens, L., Macdonald, W.H., Swanson, S.K., Glaser, R.L., Yates, J.R., 3rd, Abmayr, S.M., Washburn, M.P., and Workman, J.L. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* 306, 2084-2087.

- Lahm, A., Paolini, C., Pallaoro, M., Nardi, M.C., Jones, P., Neddermann, P., Sambucini, S., Bottomley, M.J., Lo Surdo, P., Carfi, A., *et al.* (2007). Unraveling the hidden catalytic activity of vertebrate class IIa histone deacetylases. *Proceedings of the National Academy of Sciences of the United States of America* 104, 17335-17340.
- Lalonde, M.E., Avvakumov, N., Glass, K.C., Joncas, F.H., Saksouk, N., Holliday, M., Paquet, E., Yan, K., Tong, Q., Klein, B.J., *et al.* (2013). Exchange of associated factors directs a switch in HBO1 acetyltransferase histone tail specificity. *Genes & development* 27, 2009-2024.
- Lam, K.C., Muhlpfordt, F., Vaquerizas, J.M., Raja, S.J., Holz, H., Luscombe, N.M., Manke, T., and Akhtar, A. (2012). The NSL complex regulates housekeeping genes in *Drosophila*. *PLoS genetics* 8, e1002736.
- Landry, J., Sutton, A., Tafrov, S.T., Heller, R.C., Stebbins, J., Pillus, L., and Sternglanz, R. (2000). The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proceedings of the National Academy of Sciences of the United States of America* 97, 5807-5811.
- Larschan, E., Alekseyenko, A.A., Gortchakov, A.A., Peng, S., Li, B., Yang, P., Workman, J.L., Park, P.J., and Kuroda, M.I. (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Molecular cell* 28, 121-133.
- Larschan, E., Bishop, E.P., Kharchenko, P.V., Core, L.J., Lis, J.T., Park, P.J., and Kuroda, M.I. (2011). X chromosome dosage compensation via enhanced transcriptional elongation in *Drosophila*. *Nature* 471, 115-118.
- Larschan, E., Soruco, M.M., Lee, O.K., Peng, S., Bishop, E., Chery, J., Goebel, K., Feng, J., Park, P.J., and Kuroda, M.I. (2012). Identification of chromatin-associated regulators of MSL complex targeting in *Drosophila* dosage compensation. *PLoS genetics* 8, e1002830.
- Lee, D.Y., Hayes, J.J., Pruss, D., and Wolffe, A.P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* 72, 73-84.
- Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. *Nature reviews Molecular cell biology* 8, 284-295.
- Lehmann, L., Ferrari, R., Vashisht, A.A., Wohlschlegel, J.A., Kurdistani, S.K., and Carey, M. (2012). Polycomb repressive complex 1 (PRC1) disassembles RNA polymerase II preinitiation complexes. *The Journal of biological chemistry* 287, 35784-35794.
- Li, G., Levitus, M., Bustamante, C., and Widom, J. (2005). Rapid spontaneous accessibility of nucleosomal DNA. *Nature structural & molecular biology* 12, 46-53.
- Li, X., Li, L., Pandey, R., Byun, J.S., Gardner, K., Qin, Z., and Dou, Y. (2012). The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. *Cell stem cell* 11, 163-178.
- Li, X., Wu, L., Corsa, C.A., Kunkel, S., and Dou, Y. (2009). Two mammalian MOF complexes regulate transcription activation by distinct mechanisms. *Molecular cell* 36, 290-301.

- Lim, C.K., and Kelley, R.L. (2013). The *Drosophila* over compensating males gene genetically inhibits dosage compensation in males. *PloS one* 8, e60450.
- Liu, X., Wang, L., Zhao, K., Thompson, P.R., Hwang, Y., Marmorstein, R., and Cole, P.A. (2008). The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* 451, 846-850.
- Liu, Y., Lu, C., Yang, Y., Fan, Y., Yang, R., Liu, C.F., Korolev, N., and Nordenskiöld, L. (2011). Influence of histone tails and H4 tail acetylations on nucleosome-nucleosome interactions. *Journal of molecular biology* 414, 749-764.
- Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* 49, 203-210.
- Lu, L., Li, L., Lv, X., Wu, X.S., Liu, D.P., and Liang, C.C. (2011). Modulations of hMOF autoacetylation by SIRT1 regulate hMOF recruitment and activities on the chromatin. *Cell research* 21, 1182-1195.
- Lu, X., Simon, M.D., Chodaparambil, J.V., Hansen, J.C., Shokat, K.M., and Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nature structural & molecular biology* 15, 1122-1124.
- Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nature reviews Molecular cell biology* 13, 436-447.
- Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
- Lusser, A., Urwin, D.L., and Kadonaga, J.T. (2005). Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. *Nature structural & molecular biology* 12, 160-166.
- Ma, X.J., Wu, J., Althelm, B.A., Schultz, M.C., and Grunstein, M. (1998). Deposition-related sites K5/K12 in histone H4 are not required for nucleosome deposition in yeast. *Proceedings of the National Academy of Sciences of the United States of America* 95, 6693-6698.
- Maeshima, K., Hihara, S., and Eltsov, M. (2010). Chromatin structure: does the 30-nm fibre exist in vivo? *Current opinion in cell biology* 22, 291-297.
- Maier, V.K., Chioda, M., Rhodes, D., and Becker, P.B. (2008). ACF catalyses chromatosome movements in chromatin fibres. *The EMBO journal* 27, 817-826.
- Makowski, A.M., Dutnall, R.N., and Annunziato, A.T. (2001). Effects of acetylation of histone H4 at lysines 8 and 16 on activity of the Hat1 histone acetyltransferase. *The Journal of biological chemistry* 276, 43499-43502.
- Mangenot, S., Leforestier, A., Vachette, P., Durand, D., and Livolant, F. (2002). Salt-induced conformation and interaction changes of nucleosome core particles. *Biophysical journal* 82, 345-356.
- Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature* 469, 343-349.

- McConnell, K.H., Dixon, M., and Calvi, B.R. (2012). The histone acetyltransferases CBP and Chameau integrate developmental and DNA replication programs in *Drosophila* ovarian follicle cells. *Development* 139, 3880-3890.
- Meller, V.H., Wu, K.H., Roman, G., Kuroda, M.I., and Davis, R.L. (1997). roX1 RNA paints the X chromosome of male *Drosophila* and is regulated by the dosage compensation system. *Cell* 88, 445-457.
- Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., *et al.* (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Molecular cell* 21, 811-823.
- Merson, T.D., Dixon, M.P., Collin, C., Rietze, R.L., Bartlett, P.F., Thomas, T., and Voss, A.K. (2006). The transcriptional coactivator Querkopf controls adult neurogenesis. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26, 11359-11370.
- Meyer, B.J. (2005). X-Chromosome dosage compensation. *WormBook : the online review of C elegans biology*, 1-14.
- Miescher, F. (1871). Ueber die chemische Zusammensetzung der Eiterzellen. *Hoppe-Seyler's medicinisch-chemische Untersuchungen* (1871): 441-460.
- Min, J., Feng, Q., Li, Z., Zhang, Y., and Xu, R.M. (2003). Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* 112, 711-723.
- Mingarro, I., Sendra, R., Salvador, M.L., and Franco, L. (1993). Site specificity of pea histone acetyltransferase B in vitro. *The Journal of biological chemistry* 268, 13248-13252.
- Miotto, B., Sagnier, T., Berenger, H., Bohmann, D., Pradel, J., and Graba, Y. (2006). Chameau HAT and DRpd3 HDAC function as antagonistic cofactors of JNK/AP-1-dependent transcription during *Drosophila* metamorphosis. *Genes & development* 20, 101-112.
- Miotto, B., and Struhl, K. (2010). HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by Geminin. *Molecular cell* 37, 57-66.
- Mishima, Y., Miyagi, S., Saraya, A., Negishi, M., Endoh, M., Endo, T.A., Toyoda, T., Shinga, J., Katsumoto, T., Chiba, T., *et al.* (2011). The Hbo1-Brd1/Brpf2 complex is responsible for global acetylation of H3K14 and required for fetal liver erythropoiesis. *Blood* 118, 2443-2453.
- Miyagi, A., Ando, T., and Lyubchenko, Y.L. (2011). Dynamics of nucleosomes assessed with time-lapse high-speed atomic force microscopy. *Biochemistry* 50, 7901-7908.
- Mohan, R.D., Dialynas, G., Weake, V.M., Liu, J., Martin-Brown, S., Florens, L., Washburn, M.P., Workman, J.L., and Abmayr, S.M. (2014). Loss of *Drosophila* Ataxin-7, a SAGA subunit, reduces H2B ubiquitination and leads to neural and retinal degeneration. *Genes & development* 28, 259-272.
- Moore, S.A., Ferhatoglu, Y., Jia, Y., Al-Jiab, R.A., and Scott, M.J. (2010). Structural and biochemical studies on the chromo-barrel domain of male specific lethal 3 (MSL3) reveal a binding preference for mono- or dimethyllysine 20 on histone H4. *The Journal of biological chemistry* 285, 40879-40890.

- Morales, V., Straub, T., Neumann, M.F., Mengus, G., Akhtar, A., and Becker, P.B. (2004). Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. *The EMBO journal* 23, 2258-2268.
- Moriniere, J., Rousseaux, S., Steuerwald, U., Soler-Lopez, M., Curtet, S., Vitte, A.L., Govin, J., Gaucher, J., Sadoul, K., Hart, D.J., *et al.* (2009). Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* 461, 664-668.
- Morra, R., Yokoyama, R., Ling, H., and Lucchesi, J.C. (2011). Role of the ATPase/helicase maleless (MLE) in the assembly, targeting, spreading and function of the male-specific lethal (MSL) complex of *Drosophila*. *Epigenetics & chromatin* 4, 6.
- Mueller-Planitz, F., Klinker, H., and Becker, P.B. (2013). Nucleosome sliding mechanisms: new twists in a looped history. *Nature structural & molecular biology* 20, 1026-1032.
- Mukherjee, A.S., and Beermann, W. (1965). Synthesis of ribonucleic acid by the X-chromosomes of *Drosophila melanogaster* and the problem of dosage compensation. *Nature* 207, 785-786.
- Muller, H. J., League, B. B., and Offermann, C. A.: 1931. Effects of dosage changes of sex-linked genes, and the compensatory effects of other gene differences between male and female. (Abstr.) *Anat. Rec.* 51 (Supplement): 110. 110.
- Muller, H.J., Evidence of the precision of genetic adaptation. *The Harvey Lectures*, 1950(18): p. 165–229.
- Murr, R., Loizou, J.I., Yang, Y.G., Cuenin, C., Li, H., Wang, Z.Q., and Herceg, Z. (2006). Histone acetylation by Trrap-Tip60 modulates loading of repair proteins and repair of DNA double-strand breaks. *Nature cell biology* 8, 91-99.
- Nagy, Z., and Tora, L. (2007). Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* 26, 5341-5357.
- Nelson, D.A. (1982). Histone acetylation in baker's yeast. Maintenance of the hyperacetylated configuration in log phase protoplasts. *The Journal of biological chemistry* 257, 1565-1568.
- Nguyen, A.T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. *Genes & development* 25, 1345-1358.
- Nguyen, U.T., Bittova, L., Muller, M.M., Fierz, B., David, Y., Houck-Loomis, B., Feng, V., Dann, G.P., and Muir, T.W. (2014). Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nature methods* 11, 834-840.
- Nightingale, K.P., Baumann, M., Eberharter, A., Mamais, A., Becker, P.B., and Boyes, J. (2007). Acetylation increases access of remodelling complexes to their nucleosome targets to enhance initiation of V(D)J recombination. *Nucleic acids research* 35, 6311-6321.
- Nishikori, S., Hattori, T., Fuchs, S.M., Yasui, N., Wojcik, J., Koide, A., Strahl, B.D., and Koide, S. (2012). Broad ranges of affinity and specificity of anti-histone antibodies revealed by a quantitative peptide immunoprecipitation assay. *Journal of molecular biology* 424, 391-399.

- Nishino, Y., Eltsov, M., Joti, Y., Ito, K., Takata, H., Takahashi, Y., Hihara, S., Frangakis, A.S., Imamoto, N., Ishikawa, T., *et al.* (2012). Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *The EMBO journal* *31*, 1644-1653.
- Ogiwara, H., Ui, A., Otsuka, A., Satoh, H., Yokomi, I., Nakajima, S., Yasui, A., Yokota, J., and Kohno, T. (2011). Histone acetylation by CBP and p300 at double-strand break sites facilitates SWI/SNF chromatin remodeling and the recruitment of non-homologous end joining factors. *Oncogene* *30*, 2135-2146.
- Ogryzko, V.V., Kotani, T., Zhang, X., Schiltz, R.L., Howard, T., Yang, X.J., Howard, B.H., Qin, J., and Nakatani, Y. (1998). Histone-like TAFs within the PCAF histone acetylase complex. *Cell* *94*, 35-44.
- Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). *Science* *183*, 330-332.
- Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* *4*, 281-300.
- Parthun, M.R. (2012). Histone acetyltransferase 1: More than just an enzyme? *Biochimica et biophysica acta* *1819*, 256-263.
- Parthun, M.R., Widom, J., and Gottschling, D.E. (1996). The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell* *87*, 85-94.
- Patel, D.J., and Wang, Z. (2013). Readout of epigenetic modifications. *Annual review of biochemistry* *82*, 81-118.
- Peng, J.C., and Karpen, G.H. (2009). Heterochromatic genome stability requires regulators of histone H3 K9 methylation. *PLoS genetics* *5*, e1000435.
- Pesavento, J.J., Yang, H., Kelleher, N.L., and Mizzen, C.A. (2008). Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. *Molecular and cellular biology* *28*, 468-486.
- Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., *et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Molecular cell* *12*, 1577-1589.
- Petruk, S., Sedkov, Y., Smith, S., Tillib, S., Kraevski, V., Nakamura, T., Canaani, E., Croce, C.M., and Mazo, A. (2001). Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene. *Science* *294*, 1331-1334.
- Phillips, D.M. (1963). The presence of acetyl groups of histones. *The Biochemical journal* *87*, 258-263.
- Pourfarzad, F., Aghajani-refah, A., de Boer, E., Ten Have, S., Bryn van Dijk, T., Kheradmandkia, S., Stadhouders, R., Thongjuea, S., Soler, E., Gillemans, N., *et al.* (2013). Locus-specific proteomics by TChP: targeted chromatin purification. *Cell reports* *4*, 589-600.

- Poveda, A., Pamblanco, M., Tafrov, S., Tordera, V., Sternglanz, R., and Sendra, R. (2004). Hif1 is a component of yeast histone acetyltransferase B, a complex mainly localized in the nucleus. *The Journal of biological chemistry* 279, 16033-16043.
- Qin, S., and Parthun, M.R. (2002). Histone H3 and the histone acetyltransferase Hat1p contribute to DNA double-strand break repair. *Molecular and cellular biology* 22, 8353-8365.
- Rauh, D., Fischer, F., Gertz, M., Lakshminarasimhan, M., Bergbrede, T., Aladini, F., Kambach, C., Becker, C.F., Zerweck, J., Schutkowski, M., *et al.* (2013). An acetylome peptide microarray reveals specificities and deacetylation substrates for all human sirtuin isoforms. *Nature communications* 4, 2327.
- Ravens, S., Fournier, M., Ye, T., Stierle, M., Dembele, D., Chavant, V., and Tora, L. (2014). MOF-associated complexes have overlapping and unique roles in regulating pluripotency in embryonic stem cells and during differentiation. *Elife (Cambridge)*, e02104.
- Regnard, C., Straub, T., Mitterweger, A., Dahlsveen, I.K., Fabian, V., and Becker, P.B. (2011). Global analysis of the relationship between JIL-1 kinase and transcription. *PLoS genetics* 7, e1001327.
- Reinke, H., Gregory, P.D., and Horz, W. (2001). A transient histone hyperacetylation signal marks nucleosomes for remodeling at the PHO8 promoter in vivo. *Molecular cell* 7, 529-538.
- Richman, R., Chicoine, L.G., Collini, M.P., Cook, R.G., and Allis, C.D. (1988). Micronuclei and the cytoplasm of growing *Tetrahymena* contain a histone acetylase activity which is highly specific for free histone H4. *The Journal of cell biology* 106, 1017-1026.
- Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual review of genetics* 38, 413-443.
- Robinson, P.J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R.G., and Rhodes, D. (2008). 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *Journal of molecular biology* 381, 816-825.
- Robinson, P.J., Fairall, L., Huynh, V.A., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proceedings of the National Academy of Sciences of the United States of America* 103, 6506-6511.
- Rothbart, S.B., Dickson, B.M., Ong, M.S., Krajewski, K., Houliston, S., Kireev, D.B., Arrowsmith, C.H., and Strahl, B.D. (2013). Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes & development* 27, 1288-1298.
- Rothbart, S.B., Lin, S., Britton, L.M., Krajewski, K., Keogh, M.C., Garcia, B.A., and Strahl, B.D. (2012). Poly-acetylated chromatin signatures are preferred epitopes for site-specific histone H4 acetyl antibodies. *Scientific reports* 2, 489.
- Routh, A., Sandin, S., and Rhodes, D. (2008). Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proceedings of the National Academy of Sciences of the United States of America* 105, 8872-8877.

- Rudkin, G. T. (1964). The Nucleohistones (eds Bonner, J. & Tso, P.) 184–192 (Holden-Day, San Francisco).
- Ruiz-Carrillo, A., Wangh, L.J., and Allfrey, V.G. (1975). Processing of newly synthesized histone molecules. *Science* *190*, 117-128.
- Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T.W., Patel, D.J., *et al.* (2011). Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* *145*, 692-706.
- Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nature reviews Molecular cell biology* *8*, 983-994.
- Sakabe, K., Wang, Z., and Hart, G.W. (2010). Beta-N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proceedings of the National Academy of Sciences of the United States of America* *107*, 19915-19920.
- Saksouk, N., Avvakumov, N., Champagne, K.S., Hung, T., Doyon, Y., Cayrou, C., Paquet, E., Ullah, M., Landry, A.J., Cote, V., *et al.* (2009). HBO1 HAT complexes target chromatin throughout gene coding regions via multiple PHD finger interactions with histone H3 tail. *Molecular cell* *33*, 257-265.
- Sanchez, R., Meslamani, J., and Zhou, M.M. (2014). The bromodomain: from epigenome reader to druggable target. *Biochimica et biophysica acta* *1839*, 676-685.
- Sapountzi, V., and Cote, J. (2011). MYST-family histone acetyltransferases: beyond chromatin. *Cellular and molecular life sciences : CMLS* *68*, 1147-1156.
- Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Molecular cell* *15*, 595-605.
- Sauer, M. (2013). Localization microscopy coming of age: from concepts to biological impact. *Journal of cell science* *126*, 3505-3513.
- Sauve, A.A., and Youn, D.Y. (2012). Sirtuins: NAD(+)-dependent deacetylase mechanism and regulation. *Current opinion in chemical biology* *16*, 535-543.
- Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* *436*, 138-141.
- Schermelleh, L., Heintzmann, R., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. *The Journal of cell biology* *190*, 165-175.
- Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & development* *16*, 919-932.
- Schwaiger, M., Stadler, M.B., Bell, O., Kohler, H., Oakeley, E.J., and Schubeler, D. (2009). Chromatin state marks cell-type- and gender-specific replication of the *Drosophila* genome. *Genes & development* *23*, 589-601.

- Sealy, L., and Chalkley, R. (1978). DNA associated with hyperacetylated histone is preferentially digested by DNase I. *Nucleic acids research* 5, 1863-1876.
- Segal, E., and Widom, J. (2009). What controls nucleosome positions? *Trends in genetics : TIG* 25, 335-343.
- Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. *Annual review of biochemistry* 76, 75-100.
- Sheikh, B.N., Dixon, M.P., Thomas, T., and Voss, A.K. (2012). Querkopf is a key marker of self-renewal and multipotency of adult neural stem cells. *Journal of cell science* 125, 295-309.
- Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annual review of biochemistry* 81, 65-95.
- Shinkai, Y., and Tachibana, M. (2011). H3K9 methyltransferase G9a and the related molecule GLP. *Genes & development* 25, 781-788.
- Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844-847.
- Simon, J.A., and Kingston, R.E. (2013). Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Molecular cell* 49, 808-824.
- Smith, B.C., and Denu, J.M. (2009). Chemical mechanisms of histone lysine and arginine modifications. *Biochimica et biophysica acta* 1789, 45-57.
- Smith, E.R., Allis, C.D., and Lucchesi, J.C. (2001). Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in *Drosophila* males. *The Journal of biological chemistry* 276, 31483-31486.
- Smith, J.S., Brachmann, C.B., Celic, I., Kenna, M.A., Muhammad, S., Starai, V.J., Avalos, J.L., Escalante-Semerena, J.C., Grubmeyer, C., Wolberger, C., *et al.* (2000). A phylogenetically conserved NAD⁺-dependent protein deacetylase activity in the Sir2 protein family. *Proceedings of the National Academy of Sciences of the United States of America* 97, 6658-6663.
- Sobel, R.E., Cook, R.G., Perry, C.A., Annunziato, A.T., and Allis, C.D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proceedings of the National Academy of Sciences of the United States of America* 92, 1237-1241.
- Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.M., Zhu, P., and Li, G. (2014). Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* 344, 376-380.
- Spedale, G., Timmers, H.T., and Pijnappel, W.W. (2012). ATAC-king the complexity of SAGA during evolution. *Genes & development* 26, 527-541.
- Spierer, A., Begeot, F., Spierer, P., and Delattre, M. (2008). SU(VAR)3-7 links heterochromatin and dosage compensation in *Drosophila*. *PLoS genetics* 4, e1000066.

- Spierer, A., Seum, C., Delattre, M., and Spierer, P. (2005). Loss of the modifiers of variegation Su(var)3-7 or HP1 impacts male X polytene chromosome morphology and dosage compensation. *Journal of cell science* *118*, 5047-5057.
- Sridharan, R., Gonzales-Cope, M., Chronis, C., Bonora, G., McKee, R., Huang, C., Patel, S., Lopez, D., Mishra, N., Pellegrini, M., *et al.* (2013). Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1gamma in reprogramming to pluripotency. *Nature cell biology* *15*, 872-882.
- Sterner, D.E., and Berger, S.L. (2000). Acetylation of histones and transcription-related factors. *Microbiology and molecular biology reviews* : MMBR *64*, 435-459.
- Straub, T., and Becker, P.B. (2007). Dosage compensation: the beginning and end of generalization. *Nature reviews Genetics* *8*, 47-57.
- Straub, T., and Becker, P.B. (2011). Transcription modulation chromosome-wide: universal features and principles of dosage compensation in worms and flies. *Current opinion in genetics & development* *21*, 147-153.
- Straub, T., Grimaud, C., Gilfillan, G.D., Mitterweger, A., and Becker, P.B. (2008). The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. *PLoS genetics* *4*, e1000302.
- Straub, T., Neumann, M.F., Prestel, M., Kremmer, E., Kaether, C., Haass, C., and Becker, P.B. (2005). Stable chromosomal association of MSL2 defines a dosage-compensated nuclear compartment. *Chromosoma* *114*, 352-364.
- Suganuma, T., Gutierrez, J.L., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M., and Workman, J.L. (2008). ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nature structural & molecular biology* *15*, 364-372.
- Suganuma, T., and Workman, J.L. (2011). Signals and combinatorial functions of histone modifications. *Annual review of biochemistry* *80*, 473-499.
- Sun, Y., Jiang, X., Chen, S., Fernandes, N., and Price, B.D. (2005). A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. *Proceedings of the National Academy of Sciences of the United States of America* *102*, 13182-13187.
- Sural, T.H., Peng, S., Li, B., Workman, J.L., Park, P.J., and Kuroda, M.I. (2008). The MSL3 chromodomain directs a key targeting step for dosage compensation of the Drosophila melanogaster X chromosome. *Nature structural & molecular biology* *15*, 1318-1325.
- Swigut, T., and Wysocka, J. (2007). H3K27 demethylases, at long last. *Cell* *131*, 29-32.
- Talbert, P.B., and Henikoff, S. (2010). Histone variants--ancient wrap artists of the epigenome. *Nature reviews Molecular cell biology* *11*, 264-275.
- Tamburini, B.A., and Tyler, J.K. (2005). Localized histone acetylation and deacetylation triggered by the homologous recombination pathway of double-strand DNA repair. *Molecular and cellular biology* *25*, 4903-4913.

- Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., *et al.* (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* *146*, 1016-1028.
- Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* *272*, 408-411.
- Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature structural & molecular biology* *14*, 1025-1040.
- Taylor, G., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M., and Bickmore, W.A. (2013). H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome research*.
- Tie, F., Banerjee, R., Saiakhova, A.R., Howard, B., Monteith, K.E., Scacheri, P.C., Cosgrove, M.S., and Harte, P.J. (2014). Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. *Development* *141*, 1129-1139.
- Tremethick, D.J. (2007). Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* *128*, 651-654.
- Tse, C., and Hansen, J.C. (1997). Hybrid trypsinized nucleosomal arrays: identification of multiple functional roles of the H2A/H2B and H3/H4 N-termini in chromatin fiber compaction. *Biochemistry* *36*, 11381-11388.
- Turner, B.M., Birley, A.J., and Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. *Cell* *69*, 375-384.
- Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T., and Kadonaga, J.T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* *402*, 555-560.
- Valor, L.M., Viosca, J., Lopez-Atalaya, J.P., and Barco, A. (2013). Lysine acetyltransferases CBP and p300 as therapeutic targets in cognitive and neurodegenerative disorders. *Current pharmaceutical design* *19*, 5051-5064.
- Vaquero, A., Scher, M., Erdjument-Bromage, H., Tempst, P., Serrano, L., and Reinberg, D. (2007). SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature* *450*, 440-444.
- Verdin, E., Hirschey, M.D., Finley, L.W., and Haigis, M.C. (2010). Sirtuin regulation of mitochondria: energy production, apoptosis, and signaling. *Trends in biochemical sciences* *35*, 669-675.
- Verreault, A., Kaufman, P.D., Kobayashi, R., and Stillman, B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell* *87*, 95-104.
- Verreault, A., Kaufman, P.D., Kobayashi, R., and Stillman, B. (1998). Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Current biology : CB* *8*, 96-108.

- Verzijlbergen, K.F., van Welsem, T., Sie, D., Lenstra, T.L., Turner, D.J., Holstege, F.C., Kerkhoven, R.M., and van Leeuwen, F. (2011). A barcode screen for epigenetic regulators reveals a role for the NuB4/HAT-B histone acetyltransferase complex in histone turnover. *PLoS genetics* 7, e1002284.
- Vetteese-Dadey, M., Grant, P.A., Hebbes, T.R., Crane- Robinson, C., Allis, C.D., and Workman, J.L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *The EMBO journal* 15, 2508-2518.
- Vidali, G., Boffa, L.C., Bradbury, E.M., and Allfrey, V.G. (1978). Butyrate suppression of histone deacetylation leads to accumulation of multiacetylated forms of histones H3 and H4 and increased DNase I sensitivity of the associated DNA sequences. *Proceedings of the National Academy of Sciences of the United States of America* 75, 2239-2243.
- Villa, R., Forne, I., Muller, M., Imhof, A., Straub, T., and Becker, P.B. (2012). MSL2 combines sensor and effector functions in homeostatic control of the Drosophila dosage compensation machinery. *Molecular cell* 48, 647-654.
- Voss, A.K., Vanyai, H.K., Collin, C., Dixon, M.P., McLennan, T.J., Sheikh, B.N., Scambler, P., and Thomas, T. (2012). MOZ regulates the Tbx1 locus, and Moz mutation partially phenocopies DiGeorge syndrome. *Developmental cell* 23, 652-663.
- Waldrip, Z.J., Byrum, S.D., Storey, A.J., Gao, J., Byrd, A.K., Mackintosh, S.G., Wahls, W.P., Taverna, S.D., Raney, K.D., and Tackett, A.J. (2014). A CRISPR-based approach for proteomic analysis of a single genomic locus. *Epigenetics : official journal of the DNA Methylation Society* 9.
- Wang, C.I., Alekseyenko, A.A., LeRoy, G., Elia, A.E., Gorchakov, A.A., Britton, L.M., Elledge, S.J., Kharchenko, P.V., Garcia, B.A., and Kuroda, M.I. (2013a). Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in Drosophila. *Nature structural & molecular biology* 20, 202-209.
- Wang, F., Marshall, C.B., and Ikura, M. (2013b). Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. *Cellular and molecular life sciences : CMLS*.
- Wang, L., Tang, Y., Cole, P.A., and Marmorstein, R. (2008a). Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. *Current opinion in structural biology* 18, 741-747.
- Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q., *et al.* (2008b). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature genetics* 40, 897-903.
- Watanabe, S., Radman-Livaja, M., Rando, O.J., and Peterson, C.L. (2013). A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme. *Science* 340, 195-199.
- Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature genetics* 41, 246-250.

- Widom, J., and Klug, A. (1985). Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. *Cell* 43, 207-213.
- Woodcock, C.L. (1994). Chromatin fibers observed in situ in frozen hydrated sections. Native fiber diameter is not correlated with nucleosome repeat length. *The Journal of cell biology* 125, 11-19.
- Wu, Z.Q., and Liu, X. (2008). Role for Plk1 phosphorylation of Hbo1 in regulation of replication licensing. *Proceedings of the National Academy of Sciences of the United States of America* 105, 1919-1924.
- Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nature reviews Genetics* 12, 542-553.
- Wysocki, R., Javaheri, A., Allard, S., Sha, F., Cote, J., and Kron, S.J. (2005). Role of Dot1-dependent histone H3 methylation in G1 and S phase DNA damage checkpoint functions of Rad9. *Molecular and cellular biology* 25, 8430-8443.
- Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J.D., and Zhao, Y. (2012). Lysine succinylation and lysine malonylation in histones. *Molecular & cellular proteomics : MCP* 11, 100-107.
- Xu, Y.M., Du, J.Y., and Lau, A.T. (2014). Posttranslational modifications of human histone H3: An update. *Proteomics* 14, 2047-2060.
- Yang, X., Li, L., Liang, J., Shi, L., Yang, J., Yi, X., Zhang, D., Han, X., Yu, N., and Shang, Y. (2013). Histone Acetyltransferase 1 Promotes Homologous Recombination in DNA Repair by Facilitating Histone Turnover. *The Journal of biological chemistry*.
- Yang, X.J., Ogryzko, V.V., Nishikawa, J., Howard, B.H., and Nakatani, Y. (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* 382, 319-324.
- Yang, X.J., and Seto, E. (2007). HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* 26, 5310-5318.
- Yang, X.J., and Seto, E. (2008). The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nature reviews Molecular cell biology* 9, 206-218.
- Yang, Y., Han, X., Guan, J., and Li, X. (2014). Regulation and function of histone acetyltransferase MOF. *Frontiers of medicine* 8, 79-83.
- Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A., and Korsmeyer, S.J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* 378, 505-508.
- Yuan, H., and Marmorstein, R. (2013). Histone acetyltransferases: Rising ancient counterparts to protein kinases. *Biopolymers* 99, 98-111.
- Zhang, Y., Malone, J.H., Powell, S.K., Periwai, V., Spana, E., Macalpine, D.M., and Oliver, B. (2010). Expression in aneuploid Drosophila S2 cells. *PLoS biology* 8, e1000320.
- Zhao, R., Nakamura, T., Fu, Y., Lazar, Z., and Spector, D.L. (2011). Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. *Nature cell biology* 13, 1295-1304.

- Zhao, X., Su, J., Wang, F., Liu, D., Ding, J., Yang, Y., Conaway, J.W., Conaway, R.C., Cao, L., Wu, D., *et al.* (2013). Crosstalk between NSL histone acetyltransferase and MLL/SET complexes: NSL complex functions in promoting histone H3K4 di-methylation activity by MLL/SET complexes. *PLoS genetics* 9, e1003940.
- Zheng, H., Yang, L., Peng, L., Izumi, V., Koomen, J., Seto, E., and Chen, J. (2013a). hMOF acetylation of DBC1/CCAR2 prevents binding and inhibition of SirT1. *Molecular and cellular biology* 33, 4960-4970.
- Zheng, X., Gai, X., Ding, F., Lu, Z., Tu, K., Yao, Y., and Liu, Q. (2013b). Histone acetyltransferase PCAF Up-regulated cell apoptosis in hepatocellular carcinoma via acetylating histone H4 and inactivating AKT signaling. *Molecular cancer* 12, 96.
- Zhou, J., Fan, J.Y., Rangasamy, D., and Tremethick, D.J. (2007). The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nature structural & molecular biology* 14, 1070-1076.
- Zlatanova, J., Bishop, T.C., Victor, J.M., Jackson, V., and van Holde, K. (2009). The nucleosome family: dynamic and growing. *Structure* 17, 160-171.

6 ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to my doctoral father, direct supervisor and mentor Prof. Dr. Peter Becker. The time of my PhD studies in your group was one of the most valuable experience in my life. Thank you for your continuous support, for always having an open door, for the inspiring discussions and for all the invaluable advice on scientific topics and beyond. Thank you also for providing me with the freedom to develop and test my own ideas. The training I have received from you will certainly be very influential for my future scientific and personal development.

I also want to thank the members of my thesis advisory committee Prof. Dr. Patrick Cramer, Prof. Dr. Gunter Meister, Prof. Dr. Jürg Müller, Dr. Matthias Prestel and Dr. Tobias Straub for the interest in my work, discussions and advice.

I wish to acknowledge Dr. Hans-Jörg Schäffer, Dr. Ingrid Wolf and Maxi Reif from the IMPRS coordination office for your support and superb work.

I want to thank all examiners for the fast assessment of the dissertation and Prof. Dr. Dirk Eick for writing the ‘Zweitgutachten’.

I am very grateful to had the previlige to collaborate with excellent, friendly and helpful scientists. I like to thank Dr. Johannes Söding and Dr. Holger Hartmann for the very productive collaboration, the insightful discussions and the nice working atmosphere during the weeks of my guest stay in your lab. I wish to thank Dr. Carsten Marr, Thomas Blasi and Justin Feigelman for the fruitful discussions and productive collaboration. Many thanks to the super-resolution team (Prof. Dr. Heinrich Leonhardt, Dr. Lothar Schermelleh, Dr. Jürgen Neumann and Andreas Maiser) for teaching me to use the OMX system, discussing possibilites and limitations and being always helpful and supportive! I wish to thank Dr. Ramin Norousi for the valuable work to develop software for analysing 3D-SIM data. I am indebted to Dr. Marion Cremer and Prof. Dr. Thomas Cremer for sharing expertise in FISH and many helpful and stimulating discussions on nuclear topology. I am also thankful to Dr. Pavel Tomancak for providing access to the *Drosophila* fosmid library. Also many thanks to Prof. Dr. Giacomo Cavalli and Dr. Tom Sexton for sharing data, insightful discussions and the great days I spent in Montpellier.

Thank you, Matthias, for beeing a very good, fair and open-minded teacher and for introducing and training me in fly genetics and chromatin methods. It was a pleasure to work with you and join this exciting project!

Thank you, Tobias, for the inspiring discussions, for sharing your brilliant ideas, your advice and encouragement and help with R/bioconductor.

Thank you, Ignasi, for your great help, ideas and patience during our joint project. Thank you, Ana, Ignasi and Axel, for introducing me to the world of mass spectrometers and your invaluable advice and support! Also thanks to Shahaf for the fruitful discussions and productive collaboration.

I want to thank all past and present members of the institute for creating a scientifically stimulating and friendly working atmosphere. Special thanks to all members of the Becker group: You made me enjoy every day in the lab! I'll miss the daily Mensa-talk (Dhawal, Sandro, Kenneth, Verena, Torsten, Clemens and the North lab lunch gang), the great family atmosphere in 'our' lab (Henni, Raffi, Charlotte), the teamwork and discussions (Angie, Felix, Cath, Silke, Sylvain) and the barbecues, trips, retreats etc.

Ich danke Dir, liebe Jacqueline, dass du immer für mich da bist und wir zwei so großartige Kinder haben.

Ich danke Euch, Mama, Papa und Michael, dass Ihr mich immer unterstützt habt!

7 LIST OF ABBREVIATIONS

ACF	ATP-utilizing chromatin assembly and remodelling factor
ATAC	Ada Two A containing (complex)
BPTF	bromodomain PHD finger transcription factor
CAF-1	chromatin assembly factor 1
CBP	CREB-binding protein
CHD	chromodomain-helicase-DNA-binding
ChIP	chromatin immunoprecipitation
CHRAC	chromatin accessibility complex
DCC	dosage compensation complex
DREF	DNA replication element factor
ELP3	elongation protein 3
Esa1	essential Sas2-related acetyltransferase 1
FISH	fluorescence in situ hybridization
FRAP	fluorescence recovery after photobleaching
GCN5	general control non-derepressible 5
GNAT	GCN5-related N-acetyltransferases
HAS	high affinity site
HAT	histone acetyltransferase
HBO1	histone acetyltransferase binding to ORC
HDAC	histone lysine deacetylase
HP1	heterochromatin protein 1
ISWI	imitation switch
KAT	lysine acetyltransferase
kDa	kilodalton
KDAC	lysine deacetylase
KDM	lysine demethylases

KMT	lysine methyltransferase
LC-MS	liquid chromatography mass spectrometry
MBD-R2	methyl-CpG-binding domain protein 2
MCRS2	microspherule protein 2
MLE	maleless
MLL	mixed lineage leukemia (protein)
MOF	males absent on the first
MORF	MRG-related factor
MOZ	monocytic leukemia zinc finger protein
MSL	male-specific lethal
MYST	MOZ, Ybf2/Sas3, Sas2, and TIP60
NSL	non-specific lethal
NURD	nucleosome remodelling and deacetylation
NURF	nucleosomal remodelling factor
OCM	over compensating males
OGT	O-linked β -N-acetylglucosaminetransferase
PCAF	p300/CBP associated factor
PHD	plant homeo domain
PHF20	PhD finger protein 20
PRC	polycomb repressive complex
PRMT	protein arginine methyltransferase
PTM	post-translational modification
RPD3	reduced potassium dependency 3
RNAi	RNA interference
roX	RNA on the X
RSC	remodels the structure of chromatin
SAGA	Spt–Ada–Gcn5–acetyltransferase
SAS2	something about silencing gene 2

SIR2	silent information regulator-2
SUMO	small ubiquitin-related modifier
Su(var)3-9	suppressor of variegation 3-9
SXL	sex lethal
TIP60	Tat-interacting protein 60
UAS	upstream activating sequence
WDR5	WD40 repeat protein 5
WDS	will die slowly

8 CURRICULUM VITAE

Christian Feller

Adolf Butenandt Institute, Ludwig Maximilians University Munich, Schillerstraße 44, 80336 Munich, Germany

Personal Information

Day of Birth 10th November 1983

Nationality German

Marital Status Married, two children

University Education

PhD studies with Prof. Peter Becker

Jan 2009 – present

Ludwig Maximilians University Munich, Germany

(paternity leave: Nov 2010 – Jan 2011)

Project: 'Systematic Analysis of Lysine Acetyltransferases'

Diploma thesis with Prof. Peter Becker (external)

Dec 2007 – Sep 2008

Ludwig Maximilians University Munich, Germany

Project: 'MBD-R2 as a marker for the alternative MOF complex'

Diploma studies in Biology

Oct 2003 – Sep 2008

Friedrich Schiller University of Jena, Germany

Degree: Diplom (equivalent to an MSc), Grade: 1.0 (with highest distinction)

Research Internships

Prof. William Marzluff, University of Chapel Hill, USA

Sep 2006 – May 2007

Project: 'Cloning, protein expression and initial characterization of the RNA processing factor dFip1'

Prof. Ingo Dahse, Friedrich-Schiller University of Jena, Germany

March 2006 – July 2006

Project: 'Patch-clamp studies on TTX-resistant sodium channels of DRG neurons'

Dr. Blei-Institut, Start-up company in environ. toxicology, Jena, Germany

July 2002 – July 2004

Project: 'Sample preparation and support in analytical gas chromatography'

(intermittent)

Academic Achievements

Publications

Peleg S, **Feller C**, Forné I, Schiller E, Straub T, Prestel M, Nogueira MS, Becker L, Klopstock T, Becker PB, Imhof A, Ladurner AG. Attenuation of age-associated increases in acetylation promotes *Drosophila* longevity. Submitted.

Feller C, Forné I, Imhof A, Becker PB. Global and specific responses of the histone acetylome to systematic perturbation. In Revision.

Straub T, Zabel A, Gilfillan GD, **Feller C**, Becker PB. Different chromatin interfaces of the *Drosophila* dosage compensation complex revealed by high-shear ChIP-seq. **Genome Res.** 2013 Mar;23(3):473-85.

Feller C, Prestel M, Hartmann H, Straub T, Söding J, Becker PB. The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset. **Nucleic Acids Res.** 2012 Feb; 40(4):1509-22.

Prestel M, **Feller C**, Becker PB. Genome Biol. Dosage compensation and the global re-balancing of aneuploid genomes. **Genome Biol.** 2010;11(8):216. Review.

Baumgart M, **Feller C**, Natura G, Dahse T, Schaible HG, Dahse I. Blockade of TTX-resistant and TTX-sensitive Na⁺ currents in cultured dorsal root ganglion neurons by fomocaine and the fomocaine derivative Oe 9000. **Brain Res.** 2010 Oct 28;1358:54-63.

Prestel M, **Feller C**, Straub T, Mitlöhner H, Becker PB. The activation potential of MOF is constrained for dosage compensation. **Mol Cell.** 2010 Jun 25;38(6):815-26.

Conference Talks

Keystone Conference 'Omics Meets Cell Biology', Alpbach, Austria	May 2011
'MOF associated cofactors confer different modes of transcriptional activation'	

INTERACT (Munich Life Science Student Symposium), Munich, Germany	April 2011
'MOF associated cofactors confer different modes of transcriptional activation'	

Selected Conference Posters

EMBO Conference 'Chromatin and Epigenetics', Heidelberg, Germany	June 2011
'MOF associated cofactors confer different modes of transcriptional activation'	

'Chromatin and Epigenetics' meeting, Essen, Germany	July 2010
'The activation potential of MOF is constrained for dosage compensation'	

INTERACT (Munich Life Science Student Symposium), Munich, Germany	March 2010
'The activation potential of MOF is constrained for dosage compensation'	